ATTACHMENT 1

Modeling Protocol

MODELING PROTOCOL FOR 1-HOUR OZONE MODELING IN SUPPORT OF THE SOUTH COAST AIR QUALITY MANAGEMENT DISTRICT'S YEAR 2001 AIR QUALITY MANAGEMENT PLAN UPDATE

(DRAFT #7)

July 2, 2001

Jointly Prepared by:

Planning & Technical Support Division
California Air Resources Board
P.O. Box 2815
Sacramento, CA 95812

and

South Coast Air Quality Management District
21865 E. Copley Drive
Diamond Bar, CA 91765

TABLE OF CONTENTS

ACRONYMS

Several acronyms are used in the modeling protocol document. For convenience, the acronyms used are listed below to aid the reader.

AAMA	–	American Automobile Manufacturer's Association
AGL	–	Above Ground Level
AQMP	–	Air Quality Management Plan
ARB	–	Air Resources Board
AUSPEX	–	Atmospheric Utility Signatures, Predictions, and Experiments
AVHRR	–	Advanced Very High Resolution Radiometer
BEIS	–	Biogenic Emission Inventory System
CAA	–	Clean Air Act of 1990
CAMx	–	Comprehensive Air-Quality Model with Extensions
CBM	–	Carbon Bond Mechanism
CCAA	–	California Clean Air Act
CEFS	–	California Emission Forecasting System
CMAQ	–	Community Multi-scale Air Quality (model)
CO	–	Carbon Monoxide
COG	–	Council of Governments
DARS	–	Data Attribute Rating System
DWM	–	Diagnostic Wind Model
DTIM	–	Direct Travel Impact Model
EIWG	–	Emission Inventory Working Group
EKMA	–	Empirical Kinetics Modeling Approach
EMFAC	–	Emission Factor (model)
FCM	–	Flexible Chemical Mechanism
FDDA	–	Four-Dimensional Data Assimilation
GAP	–	Geographical Approach to Protection of Biological Diversity
GCM	–	Global Climate Model
ICAPCD	–	Imperial County Air Pollution Control District
IOP	–	Intensive Operation Period
LIDAR	–	Light Detection and Ranging
MDAQMD	–	Mojave Desert Air Quality Management District
MM5	–	Mesoscale Meteorological Model (5th generation)
MWG	–	Modeling Working Group
NAAQS	–	National Ambient Air Quality Standard
NDVI	–	Normalized Difference Vegetative Index
NOx	–	Nitrogen Oxides
PDT	–	Pacific Daylight Time
PM	–	Particulate Matter
RADM	–	Regional Acid Deposition Model

RECLAIM	–	Regional Clean Air Incentives Market
ROG	–	Reactive Organic Gases
RRF	–	Relative Reduction Factor
SANDAG	–	San Diego Association of Governments
SAPRC	–	State Air Pollution Research Center
SAQM	–	SARMAP Air Quality Model
SARMAP	–	SJVAQS/AUSPEX Regional Model Adaptation Project
SBCAG	–	Santa Barbara County Association of Governments
SBCAPCD	–	Santa Barbara County Air Pollution Control District
SCAB	–	South Coast Air Basin
SCAG	–	Southern California Association of Governments
SCAQMD	–	South Coast Air Quality Management District
SCAQS	–	Southern California Air Quality Study
SCE	–	Southern California Edison
SCOS97	–	Southern California Ozone Study (1997)
SDCAPCD	–	San Diego County Air Pollution Control District
SIP	–	State Implementation Plan
SJVAQS	–	San Joaquin Valley Air Quality Study
SOx	–	Sulfur Oxides
STMPRAG	–	Scientific, Technical, and Modeling Peer Review Advisory Group
TOG	–	Total Organic Gases
UAM	–	Urban Airshed Model
USEPA	–	United States Environmental Protection Agency
UTM	–	Universal Transverse Mercator
VCAPCD	–	Ventura County Air Pollution Control District
VMT	–	Vehicle Miles Traveled
VOC	–	Volatile Organic Compound
WSPA	–	Western States Petroleum Association

INTRODUCTION

Under the federal Clean Air Act of 1990 (CAA), the South Coast Air Basin (Basin) is classified an “extreme” nonattainment area for ozone. Section 182(c)(2)(A) of the CAA sets November 15, 1994 as the deadline for submission of a State Implementation Plan (SIP) to demonstrate attainment of the federal ambient ozone air quality standard of 0.12 ppm by December 2010. The South Coast Air Quality Management District (District) satisfied that Act requirement with the submittal of the 1994 Air Quality Management Plan (AQMP) in September 1994, and a subsequent revision was submitted to the U.S. Environmental Protection Agency (USEPA) in February 1997. In order to take advantage of more recent information such as the 1997 Southern California Ozone Study (SCOS97) and enhancements to the emissions inventory, the District plans to revisit the recently submitted ozone attainment demonstration in a 2001 submittal.

The USEPA promulgated a new ozone ambient air quality standard in July 1997. While the SIP submittal date for the new federal ozone air quality standard is still being debated, the District committed to begin air quality analysis for the new standard. Subsequent to the release of the implementation guidance for the new ozone air quality standard, the U.S. Circuit Court ruled that USEPA could not enforce the new standard. USEPA is in the process of appealing that decision. However, the court decision may change the date for the SIP submittal for the new ozone air quality standard. A first look at the new ozone air quality standard was provided in the 1997 AQMP. However, the analysis was conducted before USEPA finalized the new ozone air quality standard.

The CAA requires the use of an USEPA-approved, photochemical grid model to perform the attainment demonstration. USEPA's "Guideline on Air Quality Models (Revised)" recommends the use of the Urban Airshed Model (UAM) for attainment demonstrations involving entire urban areas (USEPA, 1990). However, the USEPA recently revised its recommendation to no longer include a recommended air quality model for ozone. Instead, the USEPA recommends that air quality model(s) proposed for an ozone attainment demonstration, be subjected to model performance evaluations to demonstrate that they are appropriate for attainment demonstration purposes. The USEPA issued the "Guideline for Regulatory Applications of the Urban Airshed Model" to assist states in preparing the attainment demonstration required by the CAA (USEPA, 1991 and 1996). In addition, the ARB has issued photochemical modeling guidance for use by the districts to ensure the technical validity of the modeling results (ARB 1992). Finally, the USEPA is in the process of finalizing attainment demonstration guidance for the new, federal, 8-hour ozone air quality standards. This ozone modeling protocol is based on these guideline documents.

In order to devote the maximum resources practicable to the development of the District's 2001 AQMP, the Executive Officers of the ARB and the SCAQMD have agreed to jointly develop the emissions and air quality modeling needed to determine the carrying capacity and attainment demonstration for the 1-hour ozone and PM_{10} standards. The technical staffs of both agencies are working closely together to plan and carry out the necessary work for the AQMP, and are committed to intensive and timely coordination to ensure that the AQMP is based on the soundest science possible. Both agencies agree that their staffs will collaborate on this work such that the product will be mutually acceptable modeling analyses for use in the 2001 plan.

The objective of this protocol is to define the methodology to be used for simulating ozone formation in the basin, including: the episodes to be simulated; the model(s) to be used; the modeling domain; the model input data, including meteorology, emissions, and initial and boundary conditions; and the process for model performance evaluation. In addition, the protocol outlines the attainment demonstration process, including a review of the CCAA requirements. This protocol document is intended to be dynamic, and will be updated in response to reviewer comments and to reflect the results of new information that will emerge during the process.

Background

The first Air Quality Management Plan (AQMP) for the Basin was produced in 1979 as part of a revision to California's SIP. The 1979 AQMP indicated that it would not be possible to achieve the federal ozone air quality standard of 0.12 ppm by 1982. Because the emission controls discussed in the 1979 AQMP would not be fully effective until after 1982, the ARB and the USEPA granted an extension to 1987 for achievement of the standard. As part of that extension, a revision to the AQMP was performed by the District in 1982 which included a new series of modeling analyses to address concerns regarding the original 1979 modeling analysis.

For both the 1979 and 1982 AQMP revisions, the city-specific Empirical Kinetics Modeling Approach (EKMA) was applied. The 1979 AQMP used the city-specific EKMA procedures then in existence. The 1982 AQMP revision used a more sophisticated version of the EKMA procedures and also contained sensitivity analyses (Appendix VI-A of the 1982 AQMP revision). The UAM was used in conjunction with the EKMA analyses to evaluate the effect of applying all feasible control measures by 1987 (Appendix VI-E of the 1982 AQMP revision). On the basis of those modeling studies, it was determined that hydrocarbon reductions on the order of 75 percent or greater would be required to attain the federal standard by 1987, given a forecasted 23 percent reduction in oxides of nitrogen. Forecasted emission data indicated that only a 33 percent hydrocarbon reduction could be expected by 1987. Issues raised during the 1979 and 1982 AQMP revisions highlighted the need to use a three-dimensional, photochemical model such as the UAM to better understand the complex interactions between precursor emissions, meteorology, and the formation of ozone in the Basin.

For the 1989 AQMP revision, the UAM was applied to a single, multiday, ozone episode to demonstrate attainment of the National Ambient Air Quality Standard (NAAQS) for ozone. It was determined from the modeling analysis that hydrocarbon and oxides of nitrogen emission reductions of more than 80 percent would be needed in order to attain the NAAQS by the year 2007. The 1989 AQMP revision outlined three levels of controls (identified as Tiers I, II, and III) that separated the proposed control measures by known and proven technologies from those technologies anticipated to be available within the next 20 years.

For the 1991 AQMP, the District used the UAM to further assess the effectiveness of the three tiers of control measures in reducing ambient ozone levels. To complement the single, multiday ozone episode used for the 1989 AQMP revision, two additional ozone episodes were modeled to investigate the effect of projected emission reductions on future ozone concentrations during a wider variety of meteorological conditions. Additional evaluations of model performance, including new graphical procedures and subregional performance statistics, were used to ensure adequate representation of the physical and chemical processes that influence ozone formation in the Basin.

A number of improvements were made to the modeling analysis for the 1994 AQMP. Growth factors for population and vehicle miles traveled (VMT) were revised to reflect the 1990 Census data and the economic climate of the early 1990s, and improved transportation modeling was considered. The modeling analysis benefited from a number of District, ARB, and SCAG studies that improved the area source emission inventory (Appendix III-A). On-road, mobile emission estimates were improved with the use of the latest ARB emission factors program, EMFAC7F. Five ozone episodes were simulated to evaluate control strategy effectiveness. In addition to the June 5-7, 1985, episode used in the 1989 AQMP, and the two Southern California Air Quality Study (SCAQS) episodes (August 26-28, 1987, and June 23-25, 1987) added for the 1991 AQMP analysis, two additional episodes (July 13-15, 1987, a SCAQS episode, and September 7-9, 1987) were simulated for the 1994 AQMP. In this manner, control strategy decisions were based on a range of meteorological conditions, thereby reducing uncertainty in the control strategy's effectiveness. It was determined that hydrocarbon and oxides of nitrogen emission reductions on the order of 80 and 60 percent, respectively, would be needed in order to attain the NAAQS.

Based on the District's experience with the five ozone episodes used in preparing the 1994 AQMP, the District decided to drop the June 1985 meteorological episode for the 1997 AQMP. The District believed that the 1987 meteorological episodes were satisfactorily evaluated. Since the 1985 meteorological episode was based on routine monitored data, it was believed that the 1987 SCAQS episodes provided a greater certainty level relative to their performance evaluation. In October 1998, the District provided to the USEPA a "weight of evidence" analysis that indicated that even without the June 1985 episode, a viable ozone attainment demonstration could be made.

Since the 1997 AQMP, other ozone episodes have been observed that could serve as complementary or replacement episodes:

- As a result of intense interest for aerometric databases to support *regional* ozone modeling, a large-scale field measurement program was carried out in southern California during the Summer of 1997 to collect sufficient aerometric data to allow data analysts and modelers to characterize and simulate ozone formation and fate in the region. Several agencies and others participated during the planning and operational phases of the field study, including the ARB, the USEPA, the local districts, the US Navy, the US Marines, and the marine industry. The

1997 Southern California Ozone Study, or SCOS97, occurred over a four month period from June 15 through October 15, 1997, and captured several episodic ozone days.

- A widespread and severe ozone episode occurred throughout California in July 1998, during which many districts in southern California experienced 1-hour and 8-hour ozone violations. Fortuitously, several radar wind profilers were operating in the region at the time, providing greater than normal upper air meteorological coverage.

Overview of the Ozone Modeling Analysis

The proposed modeling analysis comprises the following tasks:

- Identify potential, new ozone meteorological episodes to be used. These episodes should represent the different meteorological conditions that are conducive to ozone formation in the Basin. The new ozone episodes would complement the 1987 episodes in the 1997 AQMP.
- Evaluate at least two state-of-the-science ozone models for the attainment demonstration, for the new episodes.
- Develop model inputs. This task includes evaluation of the raw data and of the model input files developed from them. The input files will be evaluated using graphical and other techniques.
- Simulate each episode with the proposed ozone models. This task includes a separate performance evaluation for each episode and each model. Documentation of the simulation results and performance evaluations will be provided.
- Project ozone air quality with proposed control measures in effect for the years 2007, 2010, and 2020. This task includes the required attainment demonstration. Model projections for the year 2007 are necessary since that is the year that the CAA requires attainment for severe-17 areas, such as the Coachella valley and the Mojave Desert Ozone Nonattainment Area. Ozone air quality projections to 2020 will be used to demonstrate that the control strategy maintains the federal ozone air quality standard and to establish emission budgets needed for conformity purposes.

The work to do the foregoing tasks will be divided between the District and the ARB staffs, and they will fully share all analyses, model inputs and outputs, findings, and conclusions. Consensus on each component of the analysis shall be reached before proceeding with subsequent components. In the event of technical disagreement on any of the work elements, the staffs of the District and the ARB shall attempt to reach consensus on a mutually acceptable approach. In the event that consensus cannot be reached, the disagreement will be elevated to the Executive Officers for resolution.

Schedule

Task	Due Date
1. Episode Selection	Completed
2. Air Quality and Meteorological Data Preparation	Completed
3. Emission Inventory Preparation	July 2001
4. Performance Evaluation	Late-July 2001
5. Attainment Demonstration	August 2001
6. Draft SIP Documents	August 2001
7. District Board Approval of Final SIP	November 2001
8. ARB Board Approval of Final SIP	December 2001
9. SIP Submittal to USEPA	December 2001

EPISODE SELECTION

Four ozone episodes from 1987 were simulated for the 1994 SIP and the 1997 AQMP: June 24-25, 1987; July 14-15, 1987; August 27-28, 1987; and September 8-9, 1987. To maintain continuity with recent plan submittals, the model performance for two of these episodes (June 24-25, 1987 and August 27-28, 1987) will be reevaluated using updated emission data. These two episodes have served as the controlling episodes for the 1994 and 1997 plan submittals. No new meteorological work is planned, and the updated simulations will be conducted on the SCAQS modeling domain. Because of concerns with the age of these episodes and the limited aloft data associated with them, three additional episodes from SCOS97 and 1998, as briefly described below, are proposed to supplement the SCAQS episodes.

During the four-month SCOS97 field study, the peak observed ozone concentration within the SCOS97 domain was 21 pphm. There were 13 Intensive Operation Period (IOP) days during which additional measurements were taken, such as speciated hydrocarbons and air quality aloft. The peak ozone concentration observed over all of the IOP days was 19 pphm. Because of the unique meteorological patterns during SCOS97 associated with the El Nino phenomenon, there is a concern that the peak ozone concentrations measured during SCOS97 may not represent design values for southern California. In July 1998 a severe regional ozone episode occurred in southern California which included a peak, observed, one-hour ozone concentration in the SCOS97 domain of 24 pphm. Although intensive field study data are not available for this episode, there were approximately 10 radar wind profilers operating during that period, thereby providing aloft meteorological data. These data are believed to be sufficient to support simulating an episode for this period.

Therefore, in addition to the two 1987 episodes to be simulated, three multi-day ozone episodes from SCOS97 and 1998, as discussed below, are proposed to be simulated. Synopses of the meteorology associated with each of the 1997 episodes can be found in the summary of SCOS97 field operations (ARB et. al, 1998).

1) The SCOS97 episode period of August 3-7, 1997 (Sunday–Thursday). This episode was selected because the peak, 1-hour ozone concentration of 19 pphm and the peak, 8-hour concentration of 12.5 pphm measured in

the South Coast during this period were the highest not associated with an exceptional event during SCOS97. High ozone concentrations were also observed within the Mojave Desert (1-hour peak of 14 pphm) and Ventura County (1-hour peak of 13 pphm, 8-hour peak of 11.5 pphm).

2) The SCOS97 episode period of September 26-29 (Friday–Monday) was selected because it includes the second highest, 1-hour ozone concentration measured during an IOP, and because it represents a weekend episode. The peak 1-hour (17 pphm) and 8-hour (10.7 pphm) ozone concentrations were both observed at Upland.

3) The July 13-18, 1998 (Tuesday–Saturday) episode was selected because it represents a severe, widespread high ozone event. The peak, observed, 1-hour ozone concentration in the South Coast Air Basin was 24 pphm, and the peak, 8-hour, concentration observed was 20.6 pphm. Other areas also experienced elevated ozone peaks, including San Diego (1-hour peak of 16 pphm), Ventura (1-hour peak of 17 pphm, 8-hour peak of 15 pphm), and Mojave Desert (1-hour peak of 20 pphm, 8-hour peak of 14 pphm).

Summary of Episodes to be Simulated for the Current AQMP	
Episode	Notes
• June 24-25, 1987 • August 27-28, 1987	Maintained for continuity with previous plans; updated emissions only; will utilize SCAQS modeling domain
• August 3-7, 1997	Intensive SCOS97 episode. Peak 1-hour O_3 concentration of 19 pphm; peak 8-hour O_3 concentration of 12.5 pphm in the SCAB; will utilize regional modeling domain
• September 26-29, 1997	Intensive SCOS97 weekend episode. Peak 1-hour O_3 concentration of 17 pphm; peak 8-hour O_3 concentration of 10.7 pphm in the SCAB; will utilize regional modeling domain
• July 13-18, 1998	Severe, widespread high ozone event in the region. Peak 1-hour O_3 concentration of 24 pphm; peak 8-hour O_3 concentration of 20.6 pphm in the

	SCAB; will utilize regional modeling domain

MODELING DOMAIN

As discussed previously, three new ozone episodes are planned for use in the current AQMP. These, which include two episodes captured during SCOS97 and an episode from 1998, will be simulated for a regional domain as described below. No new meteorological work is planned for the 1987 episodes, therefore they will continue to utilize the SCAQS modeling domain, which is described in previous AQMPs.

Non-SCAQS Episodes

Previous ozone modeling results in southern California, such as those in support of the 1994 State Implementation Plan (SIP), proved sensitive to boundary concentrations of air pollutants. This reflected the physical processes of recirculation of pollutants within southern California and the transport of pollutants from one air basin to another. However, because of the three-dimensional nature of transport and recirculation, it is difficult to take field study measurements that are adequate to determine boundary conditions. Thus the boundary conditions used in previous studies were uncertain and the modeling domain has been extended in an attempt to minimize their influence.

The photochemical modeling studies conducted for the 1994 SIP for the South Coast Air Basin, the San Diego Air Basin, and the South Central Coast Air Basin defined the upper domain boundary at a height of 2,000 m above ground level (AGL) or less. There were few air quality measurements above this height. However, terrain elevations in southern California often exceed 2,000 m above sea level and recirculation and transport above this height are possible. Ozonesonde measurements made during the 1997 Southern California Ozone Study (SCOS97) have shown high concentrations of ozone at heights above 3,000 m AGL.

The proposed regional modeling domain will minimize the influence of boundary conditions on simulation results and allow the effects of recirculation and interbasin transport to be better represented by meteorological and photochemical model simulations. The proposed modeling domain shown in Figure 1 will completely encompass the South Coast Air Basin and San Diego County, almost all of the South Central Coast Air Basin (excepting a small piece of San Luis Obispo County), the California-Mexico border regions, and includes most of the inland desert areas to eliminate the need to define boundary concentrations between them. The domain will extend far enough offshore to contain wind flow patterns conducive to overwater recirculation. Specifically, the UTM Zone 11 coordinates of the domain are 150-700 km UTM East and 3580-3950 km UTM North. Vertically, the modeling domain will extend to a height of at least 5,000 m AGL for a more complete representation of atmospheric processes. This will contain observed high ozone concentrations aloft and allow three-dimensional wind flow patterns near elevated terrain features to be represented better than in previous simulations, providing more accurate representation of pollutant transport and recirculation.

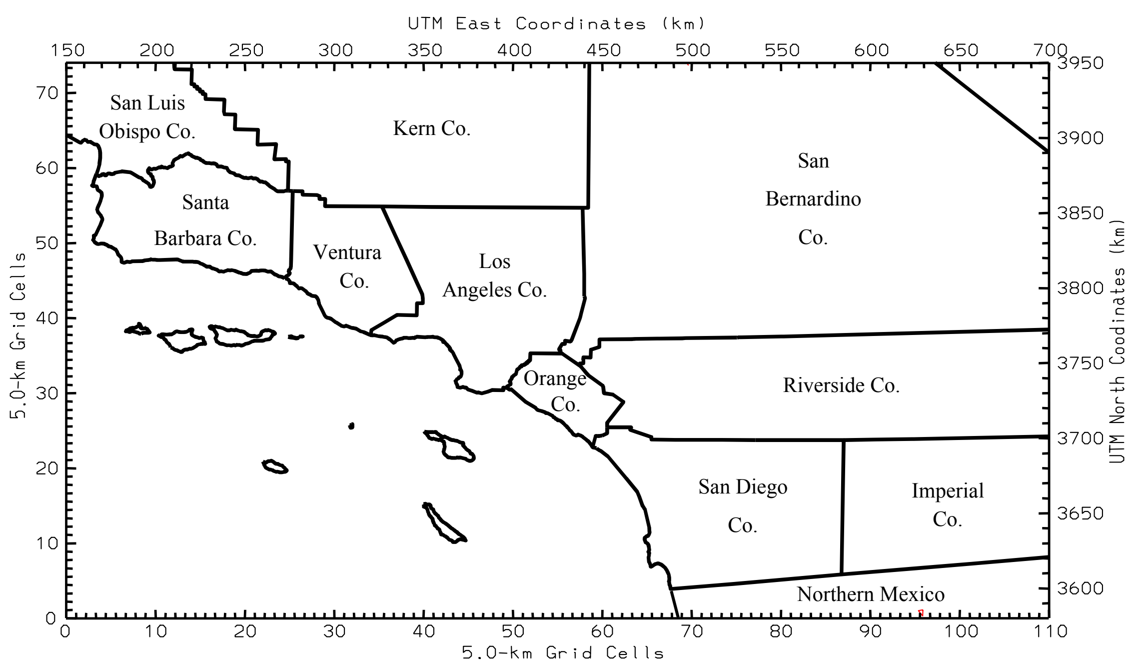


Figure 1. The regional photochemical modeling domain for the SCOS97 and the 1998 episodes.

AIR QUALITY MODEL SELECTION

The Urban Airshed Model (UAM-IV) is the only Eulerian photochemical modeling tool that has been previously approved by the USEPA for ozone modeling studies. However, the UAM-IV (USEPA 1990) is widely acknowledged to have characteristics which limit its utility when applied to large modeling domains, or to domains that are not geographically uniform. In addition, much of the science in the model is outdated, and both the USEPA and the ARB are removing that model's recommended status. A number of photochemical models have been developed over the last decade to improve upon the UAM-IV; these include:

- CALGRID (Yamartino et. al, 1989). The CALGRID model was developed for the ARB in the late 1980's. The model has been applied by various air pollution agencies around the world. It is modular to allow the user to substitute various types of wind fields and chemical mechanisms. CALGRID incorporates refined treatments of numerical advection, vertical transport and dispersion, and dry deposition. The model can be exercised with either the Carbon Bond IV (CB-IV) or SAPRC chemical mechanisms, and contains highly efficient chemical integration routines. The vertical structure of the atmosphere can be optionally defined relative to a mixing height field, similar to the UAM, or can be based on fixed layer heights and a derived mixing height.

- The Comprehensive Air-Quality Model with Extensions (Environ, 1997). CAMx contains a number of advanced features, including grid nesting, sub-grid-scale plume-in grid simulation, alternative numerical advection solvers, and the ability to use alternative chemical mechanisms. In addition it has the ability to tag emissions so that at the end of the simulation one can determine the sources of emissions impacting a particular receptor. Since CAMx is a relatively new model, there is a relatively short history of experience applying the model.

- Models-3 (USEPA, 1998a). Models-3 is a flexible software system designed for applications ranging from regulatory and policy analysis to understanding the complex interactions of atmospheric chemistry and physics. The Models-3 system is a framework that allows the user to go from developing model inputs to visualizing results all in one package. At the heart of the current version of Models-3 is the Community Multi-scale Air Quality (CMAQ) model. Its capabilities include urban to regional scale air quality simulation of ozone, acid deposition, visibility, and fine particles. CMAQ is a modular system capable of using output from the MM5 prognostic meteorological model, along with the CBIV, RADM-2, or SAPRC-99 chemical mechanism. The CMAQ model also includes a plume-in-grid module, vertical and horizontal growth due to turbulence and shear, a choice of advection schemes, and a cloud-module to simulate precipitating and non-precipitating clouds. A forthcoming aerosol module will allow simulation of $PM_{2.5}$ and PM_{10}. Since the Models-3 system is new, some implementation and application problems are likely.

- SARMAP Air Quality Model (J.S. Chang et. al, 1997). SAQM is a three-dimensional non-hydrostatic model based upon the Regional Acid Deposition Model (Chang et. al 1987, 1990). However, SAQM includes a number of improvements over RADM, including: a fixed vertical coordinate system that is compatible with MM5; a horizontal coordinate system defined in a Lambert-Conformal projection that accounts for curvature of the Earth; a mass conservation module for compatibility with non-hydrostatic meteorological inputs; the Bott advection scheme (Bott 1989a, 1989b) to reduce numerical diffusion and increase numerical accuracy; two-way nesting, and the capability to use either the CB-IV or SAPRC chemical mechanisms. A version of SAQM with plume-in-grid treatment is also available.

- Urban Airshed Model-Flexible Chemical Mechanism (Kumar et. al, 1995). The UAM-FCM is an alternate version of the UAM-IV that has been enhanced to allow the flexibility to incorporate any Carbon Bond- or SAPRC-type chemical mechanism. The FCM allows incorporation of reaction-specific photolysis rates. In addition, the UAM-FCM has a generalized methodology to solve the set of differential equations that is mechanism independent. However, the meteorological dispersion algorithms are the same as in UAM-IV.

- Urban Airshed Model-Variable (SAI, 1996). The UAM-V is an updated version of the Urban Airshed Model (UAM-IV) which incorporates many state-of-the-art enhancements in chemical mechanisms, meteorological models, and representation of emissions. Perhaps the most significant additions are: an updated CB-IV mechanism to include aqueous phase chemistry; plume-in-grid capabilities; an improved dry deposition algorithm; and an improved plume rise algorithm. Other enhancements over UAM-IV include allowing the user a fixed vertical structure as opposed to one that is relative to the diffusion break, the ability to use three dimensional inputs from prognostic models, and two-way grid nesting. However, the present non-public domain status of UAM-V may preclude regulatory usage. The model developers have indicated that the model could be made available for any party to review if the party agrees that the use of the model would be solely for the review of the AQMP.

The ozone meteorological episodes to be simulated for the current AQMP include two from 1987 for continuity with previous AQMPs. It is planned to simulate those episodes using the UAM-IV or the UAM-FCM, with updated emissions only. For the SCOS97 and later episodes, the staffs of the ARB and the SCAQMD have agreed to test at least two of the available models: CAMx and CMAQ. CAMx will be run using both diagnostic and prognostic windfields.

The models will be run using the Carbon Bond IV (Gery et. al, 1989) and SAPRC (Carter1990; Carter et. al, 1993; Carter 1995; Carter et. al, 1996; Carter et. al, 1997) chemical mechanisms. The SAPRC chemical mechanism is the basis for chemical reactivity scales.

At its meeting on October 8, 1999, the ARB's Reactivity Scientific Advisory Committee (chaired by Dr. John Seinfeld, with participation by other members Dr. Roger Atkinson, Dr. Jack Calvert, Dr. Harvey Jeffries, Dr. Jana Milford, and Dr. Armistead Russell) discussed a peer review of the SAPRC-99 mechanism conducted by Dr. William Stockwell. Members of the committee agreed that the peer review was excellent, that SAPRC-99 was a state-of-the-art chemical mechanism, and they approved the peer review. The Committee then unanimously recommended that SAPRC-99, as the most up-to-date mechanism available, be used for SIP modeling.

HORIZONTAL AND VERTICAL GRID RESOLUTION

The horizontal and vertical grid resolution for the 1987 SCAQS episodes will be unchanged; the details are provided in previous AQMPs. The remainder of this section describes the resolution for the SCOS97 and later episodes.

Horizontal Grid Resolution

Over the past decade, photochemical models have been applied in California with horizontal grid resolutions in the range from 2 x 2 km to 8 x 8 km. The specific grid resolution chosen was primarily dependent on the size of the modeling domain, computer resources available, and the time and money available to carry out the simulations. In effect the final resolution was a compromise between the accuracy desired and the cost. However, the current generation of high-speed computers have minimized cost and resource constraints.

The horizontal grid resolution plays an important role in the modeling process. Large grid resolution tends to smooth emission gradients and meteorological inputs, which in turn leads to a smoothing of the resulting concentration fields. In general, the resolution should be sufficiently small to pick up emission gradients in urban areas and be consistent with the major terrain features which may affect the air flow.

- **Air Quality Modeling**

For the year 2001 AQMP revision, a horizontal grid resolution of 5 km is proposed to be used for the air quality modeling. No grid nesting is anticipated. This resolution is consistent with the grid resolution used in earlier photochemical modeling studies for the South Coast Air Basin (e.g., SCAQMD, 1994) and for San Diego. In addition, this will reduce resources needed to create gridded emissions for the SCAQS episodes which are based on 5 km grid cells. For the proposed modeling domain, use of a 5 km resolution results in a modeling grid with 110 cells in the east-west direction, and 74 cells in the north-south direction. The CMAQ model will be based on a Lambert-Conformal map projection system, as will CAMx when used with MM5-based meteorological fields; these models will use a slightly different domain. All other air quality models will use a UTM-based horizontal coordinate system.

- **Meteorological Modeling**

Meteorological inputs to the air quality model will be provided for the same horizontal grid resolution and coordinate system (i.e., UTM or Lambert Conformal). More details of the meteorological modeling can be found in the section "METEOROLOGICAL INPUTS."

- **Emission Inventory**

The emission inventory is based on a UTM coordinate system, with a horizontal resolution of 2 km. For CMAQ, the emissions will be mapped from UTM into a Lambert Conformal coordinate system. More information on the inventory can be found in the section entitled "EMISSION INVENTORY."

Vertical Resolution

As with the selection of the horizontal grid resolution, the vertical resolution defined for photochemical modeling domains has been limited by computational resources. In addition, available aloft meteorological and air quality databases were not sufficient to characterize conditions aloft. As a result, simulation results have been limited by the relatively few number of vertical layers within the surface mixed layer, resulting in poor representation of the stratification of the atmosphere.

As enhanced aerometric databases have become available—such as the 1990 San Joaquin Valley Air Quality Study and 1997 SCOS97—more has been learned about the vertical structure of the atmosphere. The ability to better simulate the vertical structure of the atmosphere is emerging due to the availability of measurements aloft, the emergence of higher-speed computers, and our increased experience with diagnostic and prognostic meteorological models.

- **Air Quality Modeling**

To improve the vertical representation of the atmosphere, the number of vertical layers below 500 m (the nominal height of the afternoon mixing height within the Los Angeles coastal plain) will be increased from previous studies, and the modeling domain top will be set to a height of at least 5,000 m.

CAMx/MM5 – The vertical structure is shown in Table 1.

CAMx/CALMET - The fixed height vertical layers are the same as those from CALMET, and are shown in Table 2.

UAM/FCM – As for the UAM, the vertical atmosphere is defined in two zones: that above the mixing height and that below. A total of 5-8 vertical layers are proposed for the SCOS97 simulations.

CMAQ – The vertical structure is shown in Table 1.

Table 1
Proposed vertical layer heights for CMAQ and CAMx/MM5

Layer #	Height (m)*	Layer #	Height (m)*	Layer #	Height (m)*
1	58	7	737	12	1,767
2	146	8	879	13	2,094
3	250	9	1,022	14	2,942
4	369	10	1,234	15	3,962
5	490	11	1,450	16	4,986
6	613				

* *These height estimates are based on sigma-level calculations at sea level using standard conditions. Height increments will decrease as terrain elevation increases.*

- **Meteorological Modeling**

For the terrain-following CALMET model, the proposed vertical layer definition is shown below.

Table 2
Proposed vertical layer heights for CALMET/CALGRID

Layer #	Height (m)*	Layer #	Height (m)*	Layer #	Height (m)*

1................20.0 2................60.0 3...............100.0 4...............300.0 5...............400.0 6...............500.0	7 .... 600.0 8 .... 800.0 9 .... 1,000.0□ 10 .... 1,500.0□ 11 .... 2,000.0□	12 .... 2,500.0 13 .... 3,000.0 14 .... 3,500.0 15 .... 4,000.0 16 .... 5,000.0
* *Heights are for a constant-height, terrain-following coordinate system*		

For the MM5 prognostic model, the following vertical structure is proposed.

Table 3 Proposed vertical structure for MM5					
Level	Height (m)*	Level	Height (m)*	Level	Height (m)*
30	61	20	1528	10	7270
29	154	19	1862	9	7981
28	263	18	2207	8	8773
27	389	17	2714	7	9624
26	516	16	3228	6	10499
25	646	15	3837	5	11371
24	776	14	4452	4	12230
23	926	13	5083	3	13227
22	1077	12	5816	2	14334
21	1300	11	6551	1	15635

* *The vertical coordinate system for MM5 is based on a normalized pressure scale. The above layer heights were calculated from sea level using standard conditions. Layer heights are lower relative to ground level as terrain height increases.*

METEOROLOGICAL INPUTS

Air quality models require three-dimensional, meteorological inputs. The key parameters are winds, mixing heights, temperature, and insolation. The windfields describe the transport and dispersion of pollutants. Mixing heights define the vertical extent of pollutant mixing near the surface. Temperature and insolation fields influence emission rates and the rates of chemical transformation. Because meteorological measurements can be made only at discrete locations, meteorological models are required to develop the 3-dimensional fields required by models.

The meteorological models used to generate these three-dimensional fields are generally of three types: objective, diagnostic or prognostic.

Objective models are the least sophisticated meteorological models. These models rely on interpolation of observations. Obtaining a reasonable field requires sufficient observations to accurately represent the atmosphere. This is especially true for windfields. In areas with complex terrain and bodies of water, such as the proposed modeling domain, the meteorology can be quite complex, and a successful objective analysis would require an extremely large number of observations.

Diagnostic models rely both on observations and constraints based on physical concepts such as the conservation of mass. A diagnostic wind model can simulate thermally induced circulations and the effects of surface friction. One example of this type of model is the Diagnostic Wind Model (DWM) which is distributed by the USEPA. For the DWM, the user first defines an initial guess mean wind field that can be representative of synoptic scale patterns. The domain mean wind is then adjusted for the effects of terrain. Available observations are then used to develop meteorological fields using objective analysis. The initial guess and the objective analysis are then combined using a weighting function based on distance from observations. A criticism of diagnostic models is that the fields produced are not consistent from one hour to the next. Since the processes which create the wind, temperature, and mixing height fields are relatively independent, the model is criticized for not being thermodynamically consistent between the meteorological parameter fields.

Prognostic models are the most sophisticated of the meteorological models. They are based on principles of atmospheric physics, i.e., conservation of mass, momentum, energy and moisture. As a result, they are computationally intensive. The use of four dimensional data assimilation (FDDA) or observational nudging–where observations are introduced to the model as an additional forcing term–is typically used in areas of complex meteorology to improve the accuracy of the outputs. Another approach is objective combination, in which observations are introduced after the model has estimated a value. Prognostic models are capable of explicitly incorporating many of the physical flow processes important in the domain. However, prognostic models have historically had problems estimating fine-scale flow features due to the limited resolution of datasets used for describing geographic features

As indicated previously, no new meteorological work is planned for the 1987 episodes, which will be re-simulated with new emissions only for continuity with prior plans. The remainder of this section focuses on the preparation and review of meteorological inputs for the SCOS97 and later episodes.

Previous Applications

In the past, the ARB and the SCAQMD have utilized prognostic, diagnostic, and objective models to generate meteorological inputs for modeling. The National Center for Atmospheric Research's prognostic, non-hydrostatic Mesoscale Model (MM5) was applied for modeling in support of attainment planning in the San Joaquin Valley. The SCAQMD also has experience with the SAIMM prognostic model. Diagnostic models (WIND2D, WIND3D, DWM) have been applied in the Sacramento area and in southern California to prepare meteorological input fields for the application of photochemical models in those areas. The ARB and the SCAQMD have also begun a review of CALMET, which may be viewed as an improved version of the DWM and which is being distributed through the USEPA for air quality modeling applications. The CALMET model has an added feature that allows a hybrid meteorological field to be developed by merging the results from a prognostic model, such as MM5, with an objective or diagnostic analysis characteristic of the CALMET model. This hybrid approach has the potential to take advantage of the prognostic capabilities of MM5 in areas of the domain where meteorological measurements are few, and utilizing measurements in an objective analysis where there are many.

Proposed Approaches

The SCOS97 field study generated a dataset with a relatively high spatial density of meteorological observations. While this dataset suggests that an objective/diagnostic model could be adequate to develop the meteorological parameter fields required for air quality modeling, there are large portions of the modeling domain—such as over the ocean or the inland desert—where there are few observations. Therefore, three approaches are proposed to develop the necessary meteorological fields. After the fields have been developed using each approach, they will be evaluated to determine which is the most suitable for air quality modeling. This evaluation is described below.

- **Diagnostic Modeling**

The first approach will be to use the diagnostic meteorological model CALMET. As described previously, CALMET uses a fixed-height, terrain-following coordinate system. For the AQMP modeling, 16 vertical layers will be used to a height of 5,000 m above the ground (see Table 2).

- **Prognostic Modeling**

The second approach will be to use the MM5 prognostic model. The meteorological boundary conditions for MM5 are generated using the output from a Global Climate Model (GCM) with a relatively coarse grid scale of 45 km. Nested domains of 15 km and 5 km are then defined within MM5 to simulate meteorological fields for the fine grid scale of the modeling domain. The modeling domain for MM5 is defined in a Lambert-Conformal projection with two parallels to account for curvature of the Earth within the modeling domain over such a large region. Figure 2 shows the nested MM5 domains. Figure 3 shows the finest scale (interior) MM5 domain.

The vertical structure of MM5 is defined in a terrain-following, "sigma" coordinate system based upon a normalized pressure index. The 30 vertical layers defined for MM5 (see Table 3) can be transformed to fit the requirements of any air quality model.

- **Hybrid Approach**

The third approach for developing meteorological parameter fields will be to combine the results of the CALMET and MM5 models into a hybrid meteorological field. In this approach, the parameter fields will be overlaid using a weighting scheme that is based on the proximity to meteorological observations. The resultant fields benefit from the capabilities of the prognostic model in those areas of the modeling domain with few observations (such as offshore, in complex terrain, and in the desert areas), and benefit from the objective analysis component of the diagnostic model to force the fields to agree with observations. To develop the hybrid fields, the fields developed using CALMET and MM5 will need to be mapped into common horizontal and vertical coordinate domains. The CALMET model code is structured to facilitate this mapping.

Meteorological Input Validation and Technical Review

The meteorological inputs have a profound influence on the spatial and temporal resolution of ozone and other pollutant concentrations estimated by the air quality model. It is therefore essential that the products of meteorological models undergo a rigorous evaluation. By evaluating both offshore and onshore flow characteristics of the windfield and other key meteorological parameters the uncertainty in the air quality simulations can be minimized.

This process includes an evaluation of the gross circulation features in the modeling region to determine if the model is replicating those essential features (Mulberg, 1995, Lolk and Douglas, 1996). These features include areas of convergence and divergence, eddy circulations, land/sea breezes, slope flows, and transport corridors. Since the modeling domain includes large overwater areas it is also necessary to evaluate offshore flows as well.

Key features of the windfield are areas of convergence and divergence. These features result in vertical velocities which can transport pollutants upward (in the case of convergence) or bring pollutants from aloft down to the surface (with divergence). The evaluation should include a review of the convergence and divergence zones in the simulated windfield to determine if they agree with measurements or conceptual models in terms of location, timing, and extent. The impact on vertical velocities should also be evaluated.

Converging flows can sometimes result in an eddy circulation. In the SCOS97 domain two key eddy features are prevalent: the Catalina eddy (centered near Catalina Island), and the Gaviota Eddy in the Santa Barbara Channel (Smith, et. al., 1984). Both eddy circulations are important transport mechanisms; they are capable of transporting precursors and aged ozone concentrations onshore. Exceedances of the ozone standards are often observed with the presence of an eddy circulation. The timing of the onset of eddy circulation, its persistence, and spatial extent should be considered as part of the windfield validation.

Land/sea breeze circulations are another important flow feature. The sea breeze is one method whereby pollutants generated in the Los Angeles Basin are transported eastward. That is, the strength of the sea breeze will determine how far precursors and ozone generated near the coast will be transported inland. Errors in the timing of the sea breeze can cause precursor emissions to be transported to the wrong locations instead of inland where peak concentrations are observed. It is thus essential that the onset of the sea and land breezes simulated by the model be compared to observations for reasonableness.

The onshore portion of the SCOS97 domain includes areas of complex terrain. Slope flows are important as a recirculation mechanism that may influence ozone concentrations. Slope flows are probably the most challenging feature for prognostic meteorological models, due to the sparse observational data in complex terrain. A proposed qualitative approach is to determine if wind speeds estimated by the model appear to be reasonable in areas of complex terrain.

As a qualitative and quantitative evaluation of the windfields, wind speeds are proposed to be statistically summarized and plotted by site and globally throughout the SCOS97 domain (Seaman et. al., 1995, Bigler-Engler et. al., 1996). Temporal plots for key sites will be examined to determine agreement with observations. Quantitative techniques will make use of statistical measures such as the mean gross error and mean bias to compare modeled and measured wind speeds (Mulberg, 1995).

The proposed approaches for generating meteorological inputs incorporate observations, thus it should be expected that good agreement near those observation sites where data was used as input to the model. In order to diagnose the impact that incorporation of the observations has on the meteorological models, it may be useful to consider withholding some observations when executing the models to have an independent set of observations for comparison. The sites withheld should have some relation to the sites used to provide some assurance in the results from the comparison. This diagnostic evaluation is proposed to be conducted once acceptable meteorological fields have been prepared.

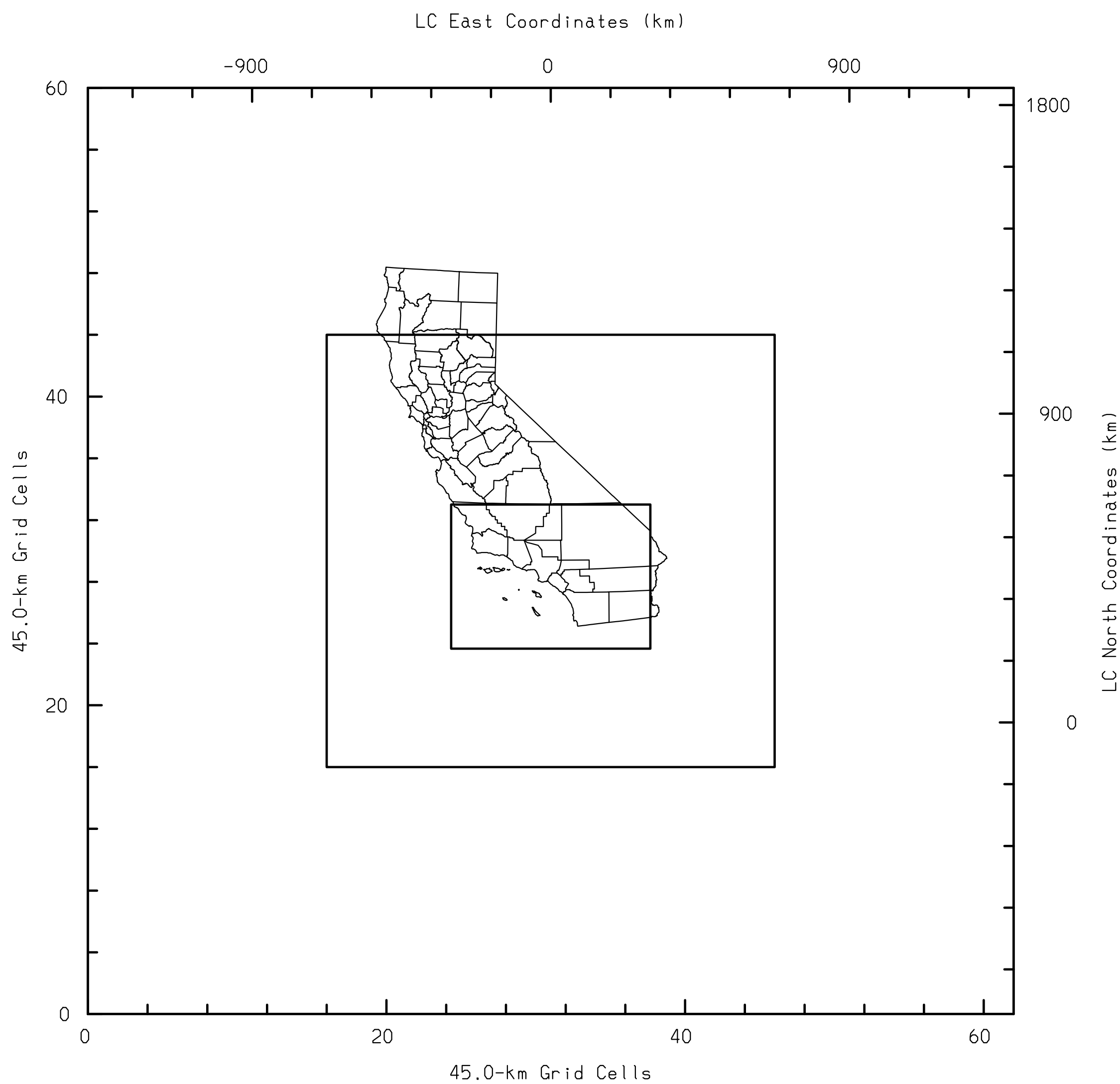
Temperature fields will also be examined. At the surface, qualitative analyses will include an examination of the temporal and spatial variation of estimated and observed temperatures. The interface at the coastline will also be examined. Mean bias and mean gross error statistics will also be calculated to provide quantitative measures of performance.

In addition, the vertical temperature profiles generated by the models will be compared to those observed at rawinsonde sites and wind profiler locations. The vertical temperature profile influences the stability characteristics of the modeling domain. One of the most notable affects is on mixing heights. The evaluation will include temporal and spatial evaluations of simulated mixing heights as compared to those estimated from observed soundings and profiler data. The timing of the onset and breakup of the inversion will also be evaluated, as this phenomenon has a profound effect on estimated ozone concentrations.

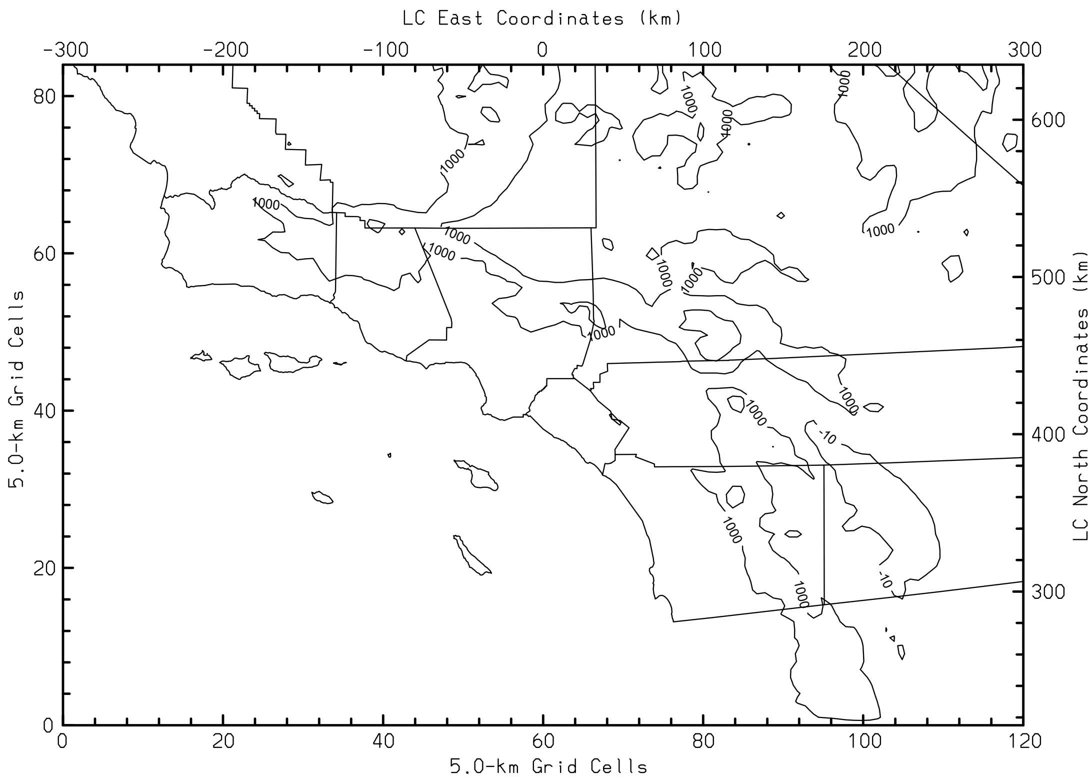
The staffs of the District and the ARB will consider the above procedures in judging the meteorological fields and in reaching consensus on the appropriateness of those fields for use in the AQMP.

<u>Figure 2</u>
Nested MM5 Domains

The horizontal grid resolution of the outermost domain is 45 km, for the middle domain is 15 km, and for the fine scale domain is 5 km.


LC East Coordinates (km)
-900
0
900
1800
900
0
LC North Coordinates (km)
60
40
20
0
45.0-km Grid Cells
0
20
40
60
45.0-km Grid Cells

<u>Figure 3</u>
The Fine-Scale (5 km) MM5 Domain.



EMISSION INVENTORY

Ozone episodes occurring in 1987, 1997, and 1998 will be simulated for the AQMP. Base year inventories for those years for CO, NOx, SOx, and TOG are needed for photochemical ozone modeling. PM is also included in the inventory in order to support inputs needed for aerosol modeling. Photochemical air quality models require gridded, hourly emissions. The information needed to complete the emission inventory for the modeling region is obtained from local air pollution control districts, transportation planning agencies, and the California Air Resources Board. In addition, contracts have been established with universities and the private sector to provide important inventory components.

1997 (SCOS97) and 1998 Episodes

The 1997 (SCOS97) and 1998 episodes will be simulated using the larger, regional modeling domain. To develop a regional emission inventory for these episodes, a SCOS97 Emission Inventory Working Group (EIWG) was formed. The EIWG is comprised of members from the ARB, SCAQMD, SDCAPCD, VCAPCD, ICAPCD, MDAQMD, SBCAPCD, and US Navy (ARB 1997a). The products of that effort, which is described below, will directly benefit the AQMP process.

Anthropogenic Emissions

- *Point Sources*

Point sources are the responsibility of the districts. Emission inventories for point sources (including RECLAIM facilities) are compiled by local districts and reported to the ARB. If annual emissions for a facility fall below 10 tons/year (this cutoff varies with district) the source is included in the area source inventory. Point sources are allocated to grid cells using the location that is stored as part of the point source emission database. Temporal codes which describe hours of operation are also included in the emission database. Factors are also stored to convert annual average emissions to a specific month and day of week.

Point sources have been inventoried for 1997. The SCAQMD's point source inventory for 1997 includes an update to locations (UTM coordinates) and stack parameters. Point source emissions for 1998 will be estimated using the ARB's California Emission Forecast System (Johnson, 1997).

- *Area Sources*

Area sources are comprised of emission source types that are difficult to inventory individually. Examples are architectural coatings, residential water heating, gasoline stations, and off-road mobile sources not included in the ARB OFFROAD model. Districts and the ARB share responsibility for estimating area source emissions according to a long-standing division of categories. Methodologies used to estimate emissions from area sources are described in several documents (ARB 1997b). Factors are also included that allow estimates of specific month and day of week emissions from annual average emissions. Temporal codes which describe hours of operation are also included in the area source emission database.

Area source categories have been inventoried for 1997. Emissions for 1998 will be grown using ARB's emission forecasting system.

- *On-Road Mobile Sources*

On-road mobile source inventories are prepared using vehicle activity data from transportation planning agencies. The majority of travel is reflected in transportation plans developed by:

- Southern California Association of Governments (SCAG);
- San Diego Association of Governments (SANDAG);
- Santa Barbara County Association of Governments (SBCAG); and
- Kern Council of Governments (Kern COG).

Travel for areas not covered by the transportation planning agencies is extracted from the California Statewide Planning Model maintained by the California Department of Transportation.

Emission factors for on-road mobile sources will be estimated using the ARB's EMFAC2000 (EMFAC2K) emission factor model. DTIM4 will use both the emission factors and travel activity data to produce hourly gridded emission estimates for the SCOS97 region.

The ARB has an extramural contract to acquire all travel data needed for this modeling study. The contractor will ensure that the digitized highways are consistent at the boundaries of the various planning areas. The network and travel activity data provided by transportation planning agencies is developed for peak and off-peak time periods, which will be processed into 24 hourly data sets. As discussed below, day-specific traffic count data will be used to calibrate DTIM4 inputs for development of day-specific on-road mobile source emissions.

ARB staff will use the network and travel activity data to produce gridded DTIM4 inventories for episode days not run by the contractor. One of the contractor's tasks is to provide training in processing inputs and executing the DTIM4 code.

- *Other Mobile Sources*

Area source emissions from most categories of off-road mobile sources will be estimated using ARB's off-road mobile source emission model (OFFROAD). OFFROAD covers more than 12 off-road categories, including lawn and garden equipment, small utility and construction equipment, as well as farm equipment. Categories not estimated by OFFROAD will be covered under "area sources". However, specific emissions for aircraft, marine vessels, and locomotives will be provided through separate special studies. OFFROAD will produce countywide emission inventories for each calendar year desired. The OFFROAD model will have the capability to estimate exhaust, starting, and evaporative emissions for differing spatial and temporal conditions.

- *Day-Specific Emissions*

Emissions from many sources vary from day to day. Evaporative emissions from vehicles and vegetation increase with ambient temperature. Exhaust emissions are also a function of ambient temperature. Increased air conditioning demands on hot days also lead to increased emissions from electrical generation. Hourly surface temperatures for episode days are interpolated to each grid cell and are used in estimating emissions from vegetation and on-road mobile sources.

Criteria pollutant emissions from approximately 80 major point sources will also be estimated hourly for each specific episode day. Each district has acquired data from major point sources for 1997 episode days and has already finished or is developing day-specific point source inventories for that year. Day-specific data for the July 1998 episode have been solicited from the districts. Districts also collected information on variances, temporary breakdowns, and shutdowns.

The results from a University of California, Davis project (that is nearing completion) to incorporate traffic count data from episode days will be used to calibrate DTIM4 inputs. The contractor will run the DTIM4 program to develop mobile source inventories for several episode days, including weekend days.

Wildfire emissions occurring on 1997 and 1998 episode days have been compiled by ARB staff, and hourly emissions have been estimated for each wildfire. The type of information collected will allow development of temporally and spatially resolved emission estimates.

A computer model to estimate emissions from large ships for the SCOS97 episodes is being developed by ARB staff. Ship activity data (for commercial vessels) from shipping ports, ship-specific engine characteristics data, and the latest emission factors will be used to estimate emissions for transit in the shipping lanes and at the ports.

The ARB has purchased aircraft activity data for the SCOS97 and July 1998 episode days. This data has one minute radar-derived locations for all aircraft in the SCOS97 region. Hourly landing, takeoff, approach, climbout, and cruise emissions can be determined from these data for each episode day. This database will be used to construct a three dimensional commercial jet emissions array that can be input to photochemical models.

In addition to the ARB efforts, the District is sponsoring special studies to collect aircraft and marine vessel activity data that will complement the ARB database.

An analysis of air quality-related special events (such as air shows, sporting events, and unusual traffic problems) has turned up an absence of such events.

Natural Emissions

- *Biogenics*

The derivation of a gridded biogenic emission inventory requires data sets describing the spatial distributions of plant species, biomass, and emission factors that define rates of hydrocarbon emissions for each plant species. The Biogenic Emission Inventory System (BEIS 2.3) (USEPA, 1995) model, distributed by the USEPA for this purpose, is one source of these data sets for areas throughout the United States. However, the BEIS model has been shown to have limited use in California because of poor spatial resolution within the referenced data sets and a simplified scheme for assigning emission factors (e.g., Jackson, et al., 1996). The development of a gridded biogenic emission inventory for the SCOS97 domain will benefit from research conducted within California that describes the needed data sets in more detail than is defined within the BEIS model (Benjamin et. al., 1998).

ARB staff in consultation with researchers at UCLA has developed a methodology to complete a gridded biogenics inventory for the SCOS97 region. The following paragraphs describe this approach.

Gridded Plant Species Maps

The distribution of plant species within the SCOS97 modeling domain will be determined using a composite of a number of data sets. Plant species as a function of urban land-use classifications were described by SCAQMD (1990) for the Los Angeles basin and updated by Benjamin et. al. (1996) and Arey et. al. (1995). These land-use classifications were extrapolated to other urban areas within the modeling domain. For the SCOS97 modeling domain, plant species distributions were taken from the GAP data base (Davis et. al., 1995), an inventory of biomass diversity for the United States.

Biomass Distribution

Plant biomass is difficult to measure and there are few descriptive data sets of biomass distribution for areas within the SCOS97 modeling domain. Therefore, estimates of biomass distribution were determined using published correlations between biomass and Normalized Difference Vegetative Index (NDVI). The NDVI is an index of relative "greeness" calculated from Advanced Very High Resolution Radiometer (AVHRR) data. Spatial distribution of NDVI from satellite remote sensing data sets were acquired from the United States Geological Survey EROS Data Center.

Emission Factors

The chemical species and rates of hydrocarbon emissions vary by plant species. Emission factors have been measured for only a relatively small subset of the more than 6,000 plant species known in California, and for only a few general categories of chemical species. However, research has shown that emission factors for various plant species can be correlated using taxonomic relationships between the plant species (Benjamin et. al., 1996). Using this “taxonomic model”, emission factors were assigned for all plant species known to exist within the SCOS97 modeling domain. However, because of the limited research results available, biogenic emission factors are available for only three classes of hydrocarbons: isoprene, monoterpenes, and methyl butenol.

Gridded Biogenic Inventory

The gridded plant species, biomass distribution, and emission factor databases will be combined with episode-specific ambient temperature and photosynthetically active radiation data using a Geographic Information System to produce gridded hourly emissions of isoprene, monoterpenes, and methylbutenol.

- *Soil NOx Emissions*

Soil emissions of NOx to the air have been associated with the use of nitrogen containing fertilizers. Emissions have been estimated from fertilizer usage in the San Joaquin Valley. The soil NOx emissions were seen to have insignificant impacts on ozone concentrations. As a result, a soil NOx inventory has not been planned for this study.

- *Oil and Gas Seeps*

There are substantial emissions from oil and gas seeps near the coast in the area of Santa Barbara county. Estimates for these sources are provided to the ARB by the SBCAPCD.

Organic Gas Speciation

Organic gas speciation profiles are applied to all categories of TOG emissions to obtain estimates for each organic gas species emitted in the modeling region. ARB maintains a database of current profiles that are routinely updated to reflect recent information. The most recent updates were for gasoline exhaust and evaporation, diesel exhaust and jet engine exhaust. These recent updates were presented by the ARB at the September 10, 1998, VOC Speciation Workshop held in Sacramento. The ARB publication "Identification of VOC Species Profiles" (ARB 1991) documents the organic gas profiles; an update is planned by summer 1999 to reflect the recent workshop.

Gridding Surrogates

Area and OFFROAD emissions are estimated and stored in the emission inventory database at the county level. There are many types of data (with highly resolved spatial resolution) that can be used as surrogates for one or more emission categories. For example, census tract population can be used to allocate consumer product emissions to grid cells. Housing units, also available for census tracts, can be used to spatially allocate residential lawn and garden equipment emissions.

Maps are used to digitize and spatially allocate emissions for several categories of watercraft, locomotives, and airports.

The ARB contractor working on the DTIM4 runs will also provide gridded surrogates for all area and off-road mobile source emission categories. The contractor will seek inputs from the districts and ARB on the appropriate gridding surrogates to use for each emission category for the SCOS97 region.

SCOS97 surrogates are being resolved to a 2 kilometer grid scale to allow high spatial resolution if needed. The 2 kilometer inventory can easily be aggregated to larger grid cells.

Northern Mexico Inventory

A portion of northern Mexico is included in the SCOS97 modeling region. A gridded inventory for this region was developed as part of a study to develop a 1990 gridded inventory for the region (SAI, 1997). The ARB contractor will also review and recommend updates to the gridded inventory for Northern Mexico. Organic gas emissions will be speciated using the most appropriate species profiles used for California emission categories.

Quality Assurance Procedures

ARB provided specific guidelines to assist state and local agencies in implementing uniform and systematic approaches for collecting, compiling, and reporting emission inventory data. A comprehensive quality control and quality assurance plan was prepared to ensure good quality practices during development of the 1997, and future year emission inventories. These procedures include: quality control checks for collecting non-emission data, updating activity data, and using appropriate emission factors for calculating emissions; emission calculation methodology; quality assurance evaluation using the Data Attribute Rating System (DARS); and quality review of the entire inventory. The DARS program, originally developed by the USEPA, will be used as an additional quality assurance tool to quantify the relative accuracy of the annual emission inventories. ARB has also provided the districts with a variety of quality assurance reports to aid in the review of inventory data important for modeling. These reports were intended to provide checks on the accuracy of the emission calculations, stack data, facility location data, temporal data, devices data, process data, etc.

Emission Projections

Future year emissions form the basis for an air quality emission reduction target. Future year emissions for area and point sources are projected by accounting for growth and control, generally using growth and control factors applied to the base year (1997) emissions. Control factors are derived based on adopted measures. Growth factors are derived from socioeconomic and demographic data provided by districts and local agencies, and ARB-sponsored research factors elsewhere. Area source and offroad emissions are gridded using the appropriate surrogates as used for 1997. The ARB contractor is also preparing gridded future year surrogates for the entire SCOS97 region for 2005, 2010, and 2020. Surrogates for other years can be interpolated as needed.

Future year traffic activity and network data are also prepared by local planning agencies. EMFAC2K will give estimates of future year emission factors. DTIM4 uses future year emission factors and network travel data to obtain gridded future year on-road mobile emissions. DTIM4 inputs for future years are being compiled and prepared (for DTIM4 input) under contract to the ARB. Ambient temperatures that occurred during 1997 are also used in calculating future year emissions for each episode day.

Biogenic emissions will not change for future years. Even though there may be a shift in farm or landscaping plans and species, the capability does not exist to incorporate any potential changes into the inventory. Seep emissions will also remain constant in future year inventories.

1987 (SCAQS) episodes

The 1987 (SCAQS) episodes will be simulated for continuity with the 1994 SIP and the 1997 AQMP, using the SCAQS modeling domain. Emissions for these episodes will be updated to reflect the following:

- a 1987 inventory re-created from a 1997 base year;
- improved VOC speciation profiles;
- ARB's EMFAC2K on-road mobile source emission factor model;
- the latest version of the Direct Travel Impact Model (DTIM4); and
- the ARB's OFFROAD (off-road) mobile source emission factor model.

In addition, work is currently underway to review stack parameter data within the SCAQMD, and the results of that effort will be incorporated when available.

INITIAL AND BOUNDARY CONDITIONS

Initial and boundary conditions for the 1987 episodes have been developed previously; no new work is planned for those episodes.

Initial Conditions

Initial conditions define the spatial distribution of chemical species concentrations throughout the 3-dimensional modeling domain at the time at which the air quality model simulation begins. There are two limitations inherent in defining initial conditions. The first is that chemical species concentrations are only measured at discrete locations and, for some species, for discrete time periods. Therefore, observed concentrations must be extrapolated to estimate concentrations throughout the modeling domain. The second limitation is that observed chemical species concentrations may not represent chemical equilibrium, especially since not all important chemical species are measured explicitly.

To minimize the importance of initial conditions on air quality model simulation results, the simulation is frequently started at some time interval before the period of interest. Historically, this "spinup" time interval has ranged between 8 and 72 hours. For the 1997 and 1998 episodes, the spinup period will begin at between 0000 PDT and 0500 PDT of the day before the first day of interest (the difference in the begin time reflects the difference in time-base – GMT vs. local time – for some meteorological models). This spinup period will allow a full diurnal cycle of sunlight, prior to the period of interest, to enable the air quality model to reach chemical equilibrium.

Initial conditions will be determined by interpolation/extrapolation of the chemical species concentration measurements available from the SCOS97 field study archive or other episode-specific data. For the large areas of the study domain in which there are few such measurements, initial-conditions will be assigned "background" values based on the minimum concentrations measured from monitoring sites where measurements are available.

Boundary Concentrations

Boundary concentrations are chemical species concentrations on the study domain boundaries and represent the concentrations for the air mass moving into the domain. Unlike initial conditions which need to be defined only for the beginning of the simulation, boundary conditions must be defined for each hour of an air quality model simulation on the 2-dimensional, vertical planes on each of the horizontal boundaries of the domain, and at the top of the modeling domain.

Ideally, modeling domain boundaries are placed so remotely that simulation results are insensitive to boundary conditions. For the SCOS97 study domain, the influence of boundary conditions on the simulation results from an air quality model is problematic. Beyond the northern SCOS97 boundary, emissions from central California could have an impact on the domain. To the south, emissions from Mexico could have an impact. The western boundary is over the Pacific Ocean, where recirculation may be an issue. Boundary conditions will be determined from measured chemical species concentrations where they are available from the SCOS97 field study. Where measurements are not available, background chemical species concentrations will be assigned based on the lowest concentrations from sites where concentrations were measured. A part of the air quality model evaluation process will be to assess the influence of boundary and initial concentrations on simulation results.

MODEL PERFORMANCE EVALUATION

It is a well established tenet of the modeling community that for an air quality modeling simulation to give reliable results, it must be capable of giving the right answers *for the right reasons*. That is, not only must the model be capable of reproducing observed ozone measurements with a reasonable level of accuracy, but it must also pass a series of prescribed tests designed to ensure that the apparently accurate results are not produced by a combination of compensating errors. As discussed below, several tests on the modeling simulations, both at the surface and aloft–for precursor and secondary species in addition to ozone–as part of the model performance process are proposed to be conducted. The statistical tests will be performed for the domain, by district, and for several subregions to ensure that the domainwide statistics do not mask subregional problems. This information should allow a determination that the model is working properly. Much of the following information is taken from the ARB's photochemical modeling guidance (ARB 1992).

Statistical and Graphical Analyses

The evaluation will include both graphical and statistical analyses. Graphical analyses compare simulated pollutant concentrations with measured values. These will include time series plots showing temporal variations, contour plots showing spatial variations, scatter plots showing tendencies for over- or underestimation, and residual plots showing the distribution of the differences between observed and predicted concentrations. The statistical analyses will examine the accuracy of peak estimates (both paired and unpaired in time and space), mean normalized bias, mean absolute gross error, and mean absolute normalized gross error. The statistical performance criteria outlined in the ARB's guidance document for Class B or better ozone performance will be used to guide the determination of acceptable model performance. These statistical criteria will be used as a criterion for acceptable model performance. However, other analyses (graphical, multi-species, aloft comparisons, and the diagnostic simulations) will also be used to determine acceptable model performance, and ultimately a conclusion that the model is working properly must be made considering all of the analyses.

- **SUBREGIONAL PERFORMANCE**

Since the South Coast Air Basin is very large, six subregions are proposed to be evaluated for model performance: San Fernando Valley, west (or coastal) basin, mid-basin, San Gabriel Valley, east basin, and the Coachella Valley area.

The same statistical acceptance criteria for the subregions as for the entire domain will be used.

Multi-Species Evaluations

To be useful for planning or other purposes, an air quality model must be able to replicate measured ozone concentrations with reasonable accuracy. However, it is also important to compare estimated and measured concentrations of ozone precursors and secondary species, to establish confidence that the chemistry processes are being simulated properly. The important precursors are NO, NO_x, HONO, and organic gas species; important secondary species are HNO_3 and PAN. Organic gas concentrations will be lumped according to the scheme employed by each model's chemical mechanism. Comparisons will be made for each of the estimated precursor species and lumped organic gas species, for each monitoring location. In addition, comparisons will also be made for NO_x, and total ROG.

These comparisons may reveal problems not associated with those for ozone concentrations. Many of the precursor species have a secondary component as well. Concentrations of primary pollutants tend to have higher gradients than do secondary species. This makes it more difficult to assume that a measured concentration of a primary pollutant represents a grid cell average. For these reasons it is probably unreasonable to expect the same accuracy in replicating precursor concentrations as for ozone concentrations. Thus use of a specific statistical error or bias criterion is not recommended. These comparisons should be viewed as qualitative, to uncover potential problems in precursor and secondary performance.

Aloft Comparisons

During the SCOS97 field study, concentrations of selected air pollutants were measured above the ground using aircraft, balloons, and LIDAR. The primary component of these measurements was oxidant concentrations measured with ozonesondes to a height of 5,000 m AGL. Ozonesondes were flown at seven sites, at 6-hour intervals, for selected episode days. Also, four aircraft were flown up to three times per day and an ozone LIDAR was operated continuously on selected episode days.

The performance of air quality model simulations above the ground will be determined by comparing simulated oxidant and ozone concentrations with measurements, for the same times and locations. Measured concentration profiles will be averaged for the vertical layer increments corresponding to those of the air quality model. Because of the vertical spacing required for the air quality models, the vertical resolution of this comparison will be relatively poor. Therefore, initially this comparison will be of a qualitative nature only.

In addition to measuring ozone, three of the aircraft measured oxides of nitrogen and collected samples for later hydrocarbon analysis. Comparisons between these precursor data and concentrations simulated using the air quality models will also be made. However, there are relatively few samples and because an aircraft is not in one grid cell for an hour, comparisons may not be consistent with modeled concentrations. Comparisons to see if any large discrepancies exist between modeled and measured concentrations aloft will be made.

<u>Diagnostic Simulations</u>

Several diagnostic, or investigative, simulations are proposed to further determine the fidelity of the model results. These are anticipated to include the following:

- *Zero emissions* – all anthropogenic and biogenic emissions will be set to zero to test the model's sensitivity to emissions and to ensure that the base case results are influenced appropriately by the emission inputs.
- *Double anthropogenic emissions* – all anthropogenic emissions will be doubled to test the model's sensitivity to increased man-made emissions. In addition, as separate tests of anthropogenic emissions affects, only mobile source emissions will be doubled and only stationary source emissions will be doubled.
- *Zero biogenics* – biogenic emissions will be set to zero to test the model's sensitivity to biogenic emissions.

- *Zero and clean air boundary and initial conditions* – the initial (interior) and boundary (at the top and sides of the modeling domain) conditions will be set to zero and USEPA recommended clean air values to determine model sensitivity to these inputs.
- *Zero surface deposition* – deposition will be turned off for all species to examine the effects of dry deposition on ozone estimations.
- *Reduced wind speeds* – reducing the wind speeds by 50% is proposed to test the model's sensitivity to that parameter. However, the feasibility of doing so in the event that a dynamically consistent, prognostic model is used, has not yet been investigated.

Acceptable Model Performance

While it is expected that acceptable model performance can be achieved for the new (SCOS97) and updated (SCAQS) ozone meteorological episodes, it is not always possible given the regulatory deadlines for plan submittals. Should the model performance of any of the new episodes not meet acceptable performance goals, that episode will not be used for carrying capacity and attainment demonstration purposes. The episode(s) will continue to be analyzed to improve on model performance for possible use in future plan revisions and SIP submittals.

USE OF THE MODELING RESULTS

We anticipate that the model results will be used for carrying capacity estimation and attainment demonstration. Use of the model results for these goals is contingent upon acceptable base case model performance for the episodes simulated. That is, only episodes for which the model is judged to be operating properly and which meet the model performance acceptance criteria will be used.

Historically, the District developed the carrying capacity and attainment demonstration for ozone based on a set of specific control measures that was projected to achieve the national 1-hour ozone air quality standard for all modeled episodes. For the 2001 AQMP revision, the ozone carrying capacity and attainment demonstration will again be based on a specific set of control measures such that the ozone concentrations predicted by the air quality model will be at or below 0.124 ppm (in order to be a viable attainment demonstration). However, we also propose to consider the use of adjustment factors (described below) to account for model bias and to use the model to address the ozone design value.

The USEPA's draft 8-hour ozone guidance (USEPA, 1998b) proposes the use of relative reduction factors (RRFs) as part of the attainment demonstration process for the 8-hour ozone standard, assuming that satisfactory base year model performance is established. The RRF approach incorporates design period monitoring data directly into the attainment test along with the ratio of future to current year model predictions. This allows the model to be used in a relative, rather than absolute, sense to reduce uncertainty in the predictions. The use of RRFs also potentially address two problems in model applications that tend to result in underestimation of emission reductions needed to attain standards. The first problem is that modeled episodes usually have ozone concentrations lower than the design value. The second problem is that simulation results have historically exhibited a tendency towards underestimation of observed concentrations. By utilizing monitored data along with model estimations, RRFs address both problems.

However, there may be some limitations in using RRFs for the one-hour standard. Examples of such situations include:

- Measured ozone concentrations at some sites and for some episodes may differ substantially from design values for those sites. That is, each available ozone episode will not be representative of design value conditions at all sites. In such instances it is unreasonable to include the non-representative sites in the RRF analysis.
- Model performance typically varies considerably between sites and episodes in a domain. The reported ozone performance measures (such as peak prediction accuracy, bias, and gross error) do not capture this variation. It may be unreasonable to include sites which have poor model performance for a given episode.

The potential use of RRFs, and details of application, will be considered after model performance evaluations are completed and analyzed. If model-estimated peaks are near design values for some episodes then RRF-type adjustments will probably not be necessary.

Carrying Capacity Estimation

A traditional use of models for planning has been the estimation of carrying capacities for ozone precursors. This is typically achieved by exercising the model with a series of across-the-board precursor emission reductions from the future year baseline, from which an ozone isopleth ("EKMA") diagram is constructed. The metric used for the isopleth diagram can be one of several, such as peak 1-hour or 8-hour ozone concentrations within the modeling domain (or subregion), number of grid cells above the standard, or one of many population exposure metrics. *Since the carrying capacity for each precursor is based on across-the-board emission changes, rather than source- and location-specific controls as would be specified in a plan, it should be viewed as an initial estimate only.*

For the AQMP process, ozone isopleth diagrams by episode for the following air quality metrics will be constructed:

- Peak 1-hour ozone concentration for the domain.
- Population exposure for 1-hour ozone concentrations.
- Peak 8-hour ozone concentration for the domain. This information will serve as an indicator of the need for potential additional precursor emission reductions to meet the 8-hour ozone NAAQS.

The resulting isopleth diagrams will be used as an initial guide for determining the emission reductions necessary for attainment.

Attainment Demonstration

The attainment strategy will be developed as in the past. All proposed control measures (the control strategy) will be modeled with the future year (2010) emission inventory to predict if the control strategy achieves the ozone standard for the episode(s) simulated. When predicted ozone concentrations in all grid cells are 124 ppb or less for each episode this step is completed.

Depending on the results from model performance evaluations, the use of RRFs will be considered, to account for differences between the episodes simulated and the ozone design values.

TECHNICAL OVERSIGHT

The District has established a technical oversight committee to review the technical aspects of the ongoing modeling analyses. During the adoption of the 1989 AQMP revision, the District's Governing Board passed a resolution to form such a group to provide oversight and to enhance technical consensus on AQMP modeling issues. The District has budgeted for the formation of a Modeling Working Group (MWG) comprised of individuals with photochemical and aerosol modeling expertise.

In 1997, the District Governing Board reconstituted the MWG and formed the Scientific, Technical, and Modeling Peer Review Advisory Group (STMPRAG). The STMPRAG role expands upon that of the MWG and includes experts in socioeconomic assessment and human health.

The STMPRAG assists the District in resolving technical issues related to air quality modeling by providing ongoing technical review of procedures and analyses. The objectives of the STMPRAG are as follows:

1. Suggest methods to gather and process meteorological, aerometric, and emission data with a specific focus on air quality modeling.

2. Provide technical guidance to the District's air quality modeling efforts, with an emphasis on ozone and PM_{10}. Areas of technical guidance include:

    - Formulation of modeling approaches;
    - Selection and development of appropriate modeling techniques; and
    - Identification of model performance evaluation methods.

3. Review and provide comments on the District's modeling procedures and analysis.
4. Make recommendations on future resource requirements (i.e., staffing and computational needs).
5. Recommend methods for interpretation of modeling results.

The MWG consists of representatives from the District, ARB, USEPA, American Automobile Manufacturers Association (AAMA), Southern

California Association of Governments (SCAG), Southern California Edison (SCE), Western States Petroleum Association (WSPA), and special technical consultants.

In addition to the STMPRAG, many of the technical work elements are being separately reviewed by a series of technical working groups established as part of the SCOS97 regional modeling effort. An Emission Inventory Working Group has already been established, and has met several times to discuss inventory-related issues and coordinate inventory preparation activities. A Meteorological Modeling Working Group has been convened with the goal of reviewing the preparation of regional meteorological inputs. Other working groups are planned, such as for air quality modeling. The participation of these groups will provide a valuable additional source of technical review to that of the STMPRAG.

Finally, as progress is made and products are available, interim results will be shared with the interested public at appropriate times and locations.

REFERENCES

Arey, J., D.E. Crowley, M. Crowley, M. Resketo, and J. Lester: 1995. "Hydrocarbon emissions from natural vegetation in California's South Coast Air Basin." Atmos. Environ. 21: 2977-2988.

Benjamin, M.T., M. Sudol, L. Bloch, and A.M. Winer. 1996. "Low-emitting urban forests: A Taxonomic methodology for assigning isoprene and monoterpene emission rates." Atmos. Environ. 30: 1437-1452.

Benjamin, M.T., A.M. Winer, J. Karlik, S. Campbell, B. Jackson, and A. Lashgari. 1998. "Assembling a biogenic hydrocarbon emissions inventory for the SCOS97-NARSTO modeling domain." Proceedings of the A&WMA 91st annual Meeting. Paper No. 98-WP75.08. A&WMA. Pittsburg, PA 15222.

Bigler-Engler, V., H. Brown, and K. Wagner, 1996. "A Hybrid Modeling Technique to Address Coastal Meteorology and Complex Terrain in the San Diego Photochemical Modeling Domain." Proceedings Ninth Joint Conference on Applications of Air Pollution Meteorology with A&WMA, Atlanta.

Bott, A. 1989a. "A Positive Definite Advection Scheme Obtained by Non-Linear Renormalization of the Advective Fluxes." Mon. Wea. Rev., 117:1006-1015.

Bott, A. 1989b. "Reply." Mon. Wea. Rev., 117:2633-2636.

California Air Resources Board. 1991. Identification of Volatile Organic Compound Species Profiles. August 1991.

California Air Resources Board: 1992. TECHNICAL GUIDANCE DOCUMENT: Photochemical Modeling. April, 1992.

California Air Resources Board. 1997a. The 1997 Southern California Ozone Study–NARSTO: Preparation of the 1997 Gridded Emission Inventory. A&WMA June 1998 Presentation.

California Air Resources Board. 1997b. Emission Inventory Procedures Manual: Methods for Assessing Area Source Emissions, Volume III. October 1997.

California Air Resources Board, and the SCOS97-NARSTO Technical Committee. 1998. SCOS97-NARSTO 1997 Southern California Ozone Study and Aerosol Study, Volume III: Summary of Field Operations. April 1998.

Carter, W.P.L. 1990. "A Detailed Mechanism for the Gas-Phase Atmospheric Reactions of Organic Compounds." Atmos. Environ. 24A:481-518.

Carter, W.P.L., D. Luo, I.L. Malkina, and J.A. Pierce. 1993. An Experimental and Modeling Study of the Photochemical Ozone Reactivity of Acetone. Final Report to the Chemical Manufacturers Association. Contract No. KET-ACE-CRC 2.0. December 10, 1993.

Carter, W.P.L. 1995. "Computer Modeling of Environmental Chamber Studies of Maximum Incremental Reactivities of Volatile Organic Compounds." Atmos. Environ., 29:2513-2527.

Carter, W.P.L., D. Luo, and I.L. Malkina. 1996. Environmental Chamber Studies for Development of An Updated Photochemical Mechanism for VOC Reactivity Assessment. December, 1996. Prepared by the University of California, Riverside for the California Air Resources Board.

Carter, W.P.L., D. Luo, and I.L. Malkina. 1997. Environmental Chamber Studies for Development of An Updated Photochemical Mechanism for VOC Reactivity Assessment. Final Report to the California Air resources Board, Coordinating Research Council, and National Renewable Energy Laboratory. November 26.

Chang, J.S., R.A. Brost, I.S.A. Isaksen, S. Madronich, P. Middleton, W.R. Stockwell, and C.J. Walcek. 1987. "A Three-Dimensional Eulerian Acid Deposition Model. Physical Concepts and Formulation." J. Geophys. Res., 92:14681-14700.

Chang, J.S., P. Middleton, W.R. Stockwell, C.J. Walcek, J.E. Pleim, and H.H. Lansford. 1990. The Regional Acid Deposition Model and Engineering Model. Final Assessment Reports of the National Acid Precipitation Assessment Program. V. 4. Government Printing Office, Washington, D.C.

Chang, J.S., S. Jin, Y. Li, M. Beauharnois, C.L. Lu, and H. Huang. 1997. The SARMAP Air Quality Model. April, 1997. Prepared by the Atmospheric Sciences Research Center, State University of New York.

Davis, F.W., P.A. Stine, D.M. Stoms, M. I. Borchert, and A. D. Hollander. 1995. "GAP analysis of the actual vegetation of California 1. The southwestern region." Madrono. 42: 40-78.

ENVIRON. 1997. User's Guide to the Comprehensive Air Quality Model with Extensions. April, 1997.

Gery, M.W., G.Z. Whitten, J.P. Killus, and M.C. Dodge. 1989. "A Photochemical Kinetics Mechanism for Urban and Regional Scale Computer Modeling." J. of Geophys. Res. 94: 12,925-12,956.

Jackson, B.S., E. Mulberg, and N. Wheeler. 1996. "The application of biogenic emission inventory estimates to photochemical modeling in California." Proceedings of the 9th Joint Conference of the American Meteorological Society and A&WMA. January, 1996. Pages 575-579. American Meteorological Society. Boston, MA. 02108.

Johnson, M. 1997. Redesign of California's Emission Forecasting System (CEFS). California Air Resources Board. A&WMA Presentation October 1997 presentation.

Kumar, N., F.W. Lurmann, and W.P.L. Carter. 1995. Development of the Flexible Chemical Mechanism Version of the Urban Airshed Model. August, 1995. Prepared by Sonoma Technology, Inc. for the California Air Resources Board.

Lolk, N.K. and S.G. Douglas. 1996. "Evaluation of Meteorological Fields Generated by a Prognostic Mesoscale Model Using Data Collected During the 1993 GMAQS/Coast Field Study." Proceedings Ninth Joint

Conference on Applications of Air Pollution Meteorology with A&WMA, Atlanta.

Mulberg, E. 1995. "A Comparison of Three Wind Models for the Broader Sacramento Area." Proceedings Regional Photochemical Measurement and Modeling Studies, A&WMA, San Diego.

SCAQMD. 1990. "Inventory of Leaf Biomass and Emission Factors for Vegetation in the South Coast Bain Basin." Technical Report III-C. Draft Air Quality Managenent Plan. South Coast Air Quality Management District. Diamond Bar, CA. 91765.

SCAQMD. 1994. "Ozone Modeling - Performance Evaluation." Technical Report V-B. Air Quality Managenent Plan. South Coast Air Quality Management District. Diamond Bar, CA. 91765.

Seaman, N.L., D.R. Stauffer, and A.M. Lario. 1995. "A Multi-Scale Four Dimensional Data Assimilation System Applied in the San Joaquin Valley During SARMAP. Part 1: Modeling Design and Basic Performance Characteristics." Journal of Applied Meteorology, 34:1739-1776.

Sonoma Technology, Inc. 1996. "Development of a Gridded Leaf Biomass Inventory for use in Estimating Biogenic Emissions for Urban Airshed Modeling". Final Report STI-996086-1599-R. Ventura County APCD. Ventura, CA. 93003-5401. August 28, 1996.

Smith, T.B., W.D. Saunders, and D.M. Takeuchi. 1984. Application of Climatological Analysis to Minimize Air Pollution Impacts in California. Prepared for the California Air Resources Board, Sacramento. Report A2-119-32.

Systems Applications, Intl. 1996. User's Guide to the Variable-Grid Urban Airshed Model (UAM-V). October 1996. SYSAPP-96-95/27r.

Systems Applications, Intl. 1997. Preparation of a Draft 1990 Gridded Emission Inventory for Southern California. March 1997. Prepared for the California Air Resources Board. SYSAPP-97/08.

United States Environmental Protection Agency. 1990. User's Guide for the Urban Airshed Model, Volume I: User's Manual for UAM (CB-IV). Office of Air Quality Planning and Standards. EPA-450/4-90-007A.

United States Environmental Protection Agency. 1991. Guideline For Regulatory Application of the Urban Airshed Model. EPA Publication No. EPA-450/4-91-013. U.S. Environmental Protection Agency, Research Triangle Park, NC.

United States Environmental Protection Agency. 1995. Urban Airshed Model (UAM) Biogenic Emission Inventory System Version 2. (BEIS2) User's Guide. Final Report. EPA Contract No. 68-D3-0034. Work Assignment No. 1-9. EC/R Project No. AQM-108. United States Enviromental Protection Agency. Air Quality Modeling Group. RTP, NC 27711.

United States Environmental Protection Agency. 1998a. EPA Third-Generation Air Quality Modeling System: Models-3 Volume 9b User Manual. June 1998. Office of Research and Development. EPA-600/R-98/069(b).

United States Environmental Protection Agency. 1998b. Use of Models and Other Analyses in Attainment Demonstrations for the 8-Hour Ozone NAAQS (Draft). October, 1998.

Yamartino, R.J., J.S. Scire, S.R. Hanna, G.R. Carmichael, and Y.S. Chang. 1989. CALGRID: A Mesoscale Photochemical Grid Model, Volume I: Model Formulation Document. June, 1989. Prepared for the California Air Resources Board by Sigma Research Corp.

ATTACHMENT 2

Expert Panel Modeling Critiques

Final Draft

Peer Review of Draft 2003 Documentation of Air Quality Modeling Used in Demonstrating Attainment of the 1-Hour federal Ozone Air Quality Standard

Report to the

South Coast Air Quality Management District
Attn: Joe Cassmassi
21865 E. Copely Drive
Diamond Bar, CA 91765-4182
Contract No. 03241

January 28, 2003

C. Shepherd Burton, Ph. D.
Consultant
2047 Huckleberry Road
San Rafael, CA 94903

Peer Review of Draft 2003 AQMP Documentation of the Air Quality Modeling Used in Demonstrating Attainment of the 1-Hour Federal Ozone Air Quality Standard

Background

The South Coast Air Quality Management District, while concluding its preparation of the draft 2003 Air Quality Management Plan (AQMP), commissioned this review, which emphasizes ozone modeling. The 2003 Plan will propose revisions to the policies and measures that restrain air emissions throughout the South Coast Air Basin. The purpose of the restraints, in the case of ozone, is to provide for attainment of the federal 1-hour ozone standard by 2010. The 2003 Plan, "as with previous updates, will employ the most recent technical information and best available science to shape the policy choices defining the optimal path to clean air." (www.aqmd.gov/aqmp/docs/Preview2003AQMP.pdf, p.44)

In the course of air quality planning, uncertainties – risks – multiply: the search for the path to clean air can seem chaotic, feel shaky and appear murky. Beginning with the 1989 AQMP, the path to clean air has been paved with some stones shaped by the UAM (Urban Airshed Model) modeling system. The system helps to order the process, provide a scientific footing and to sharpen the distinction between science and policy – the right stones seem to be in the right place. The recorded reductions in ozone concentrations across the Basin suggest the process and the contribution from modeling is working.

Air quality modeling is essential to the planning process: it is the only reasonably credible means to estimate the effects of emission changes on ozone, which is not emitted directly into the atmosphere but instead formed from a complex series of chemical reactions of oxides of nitrogen in the presence o f volatile organic compounds and sunlight. Ozone will reach unhealthful concentrations in the Basin when sufficient quantities of NO_x and VOC gases are released to the atmosphere during meteorological conditions that favor the accumulation of pollutants rather than their ventilation. Modeling ozone formation for the meteorological conditions that are conducive to ozone formation is, then, the only available means to judge the likelihood that planned policies and measures to manage the emission precursors of ozone in the Basin will provide for attainment of the ozone standard.

Air quality modeling – its science and practice – is advancing. Now there are more models and modeling methods – more tools – than there were during preparation of the 1989 Plan. These advances provide hope that the credibility of air quality management plans may be improved – that, by using better modeling, the uncertainties that increase the risks of either not achieving healthy air or incurring unnecessarily high costs for achieving it may be reduced.

AQMD and CARB technical staff, in their first formal AQMP-modeling collaboration, has expanded its modeling tools. (CARB & SCAQMD, 2001). The technical staff planned to take advantage of new data from the 1997 Southern California Ozone Study (SCOS97) and new tools from the community of model developers, and specifically to select new meteorological conditions to model, to compare different methods for preparing meteorological inputs for modeling, to compare the results obtained using several models and to select a “best” model (modeling system) based on its ability to reproduce the observations for the periods being modeled.

The model competition would be fair because all models would be run using the same geographic region, grid structure*, initial and boundary conditions and emissions inventory. The best model would be used for the 2003 AQM Planning work, in conjunction with the previously used episodes (meteorological conditions that are conducive to ozone formation) and previously used modeling methods to assure continuity with previous AQMPs. Their aim, then, was to assure that the promise from better theory, as implemented in new models, methods and procedures, was objectively and manifestly achieved – that the credibility of the 2003 Plan would exceed previous plans, that the selection and application of these new models and episodes would reflect the best application of the current knowledge and science.

The scope and plan of work was unique in California – no previous AQMP had considered multiple models. It was also ambitious: at least two of six candidate transport models (CAM_x and CMAQ), three distinct approaches for preparing meteorological inputs for five distinct modeling periods (episodes) and two alternative chemical mechanisms (CB-IV and SAPRCC-99) were considered, for a total of 60 cases (=2 x 3 x 5 x 2) to compare and evaluate.** In effect, 12 alternative modeling systems (=2 transport models x 3 meteorological input methods x 2 chemical mechanisms) would be compared, each defined by its formulation of pollutant movement, its meteorological input and its method for describing the chemistry of ozone formation. Moreover, since two distinctly different chemical mechanisms would be used, two VOC emission inventories would be processed for modeling. Finally, while the basis for this ambitious modeling endeavor was decisive in the belief that newer models would produce better results, the endeavor was indecisive about how to select a “best”

* Models of ozone formation are called grid models (or photochemical grid models) because the mathematical equations that describe the formation and transport of ozone and ozone precursors are solved using a 3-dimensional grid structure for the geographical region of interest.

** The six transport models considered with a rough indication of their date of publication were: CALGRID (1989), CAMx (1997), CMAQ (1998), SARMAP (1997), UAM-FCM (1995) and UAM-V (1995). These six together with UAM6.21, the model of record for prior AQMPs, bring the total number of models to seven.

(performing) model – the model that best reproduced the observations? – from the 12 alternative (or competing) systems.

Like many ambitious endeavors, work seldom proceeds precisely according to plan: the unexpected causes plans to change. Perhaps delays in receiving/processing episode-specific emissions, inexperience with CAM_x and CMAQ as well as the inability of the CAM_x and CMAQ models to reliably reproduce the recorded ozone concentrations for all five episodes contributed to reducing the scope of the original plan.

By a process that is undocumented, five models are now being considered for the 2003 AQMP: UAM6.21, CALGRID-CB-IV, CALGRID-SAPRC-99, CAMx-CB-IV, and CAMx-SAPRC-99. Each of these five models has been used to calculate ozone formation for the August 3 – 5, 1997 episode. Each also has been used to estimate the emissions carrying capacity[***] of the Basin.

Scope and Method of Review

AQMD requested completion of the following three tasks in this review:

1. "Review [of] the Draft AQMP documentations of the air quality modeling used in demonstrating attainment of the federal ozone and/or particulate matter air quality standards;
2. "Meet with AQMD staff, as needed, to obtain additional information and provide preliminary assessment on the AQMP modeling;
3. "Provide a final critique of the Draft 2003 AQMP air quality modeling."

In performing the review, AQMD also required that the review "consider the following questions, given the available meteorological, ambient air quality and emissions data, and the regulatory schedule for the 2003 AQMP development:

1. "Do the air quality modeling approaches used in the AQMP follow modeling practices conducted by other scientific and regulatory organizations?
2. "Does the air quality modeling follow regulatory guidelines for application of air quality models?
3. "I[s] there sufficient information to support the attainment demonstration provided in the Draft AQMP?
4. "What recommendations do you have to improve or enhance future AQMP air quality modeling efforts?"

This review draws on considerable amounts of material provided by SCAQMD. It fills two three-inch binders and five one-inch binders, and includes the most recent version of Modeling Protocol (July 2001, Draft #7), five presentations to the Scientific, Technical, Modeling, Peer Review Committee Presentations

[***] Emissions carrying capacity is an estimate, obtained by modeling, of the amount of emissions that an airshed can hold without the formation of unhealthy concentrations of ozone.

(2001- 2002), a description of current air quality in the Basin (Appendix II of Preliminary Draft 2003 AQMP), tabulations of base and future year emission inventories (Appendix III of Preliminary Draft 2003 AQMP) as well as graphical, tabular and statistical summaries of observed and calculated pollutant concentrations for each of the five modeling systems.

The review also considers EPA's current regulations for modeling that is used in air quality analyses relative to State Implementation Plans. While EPA proposed revisions to these regulations in April 2001, that proposal has not been made final; the July 1999 edition of Appendix W to Part 51-Guideline on Air Quality Models remains applicable, at least formally.

There are items not addressed in this review. For several reasons, this review neither draws inferences nor speculates about the modeling results obtained with the different modeling systems. First, AQMD did not provide model inputs, and had they provided them, the time would have been too short for their consideration. Second, chemical mechanisms with identical names may differ because, when implemented in a transport model, modelers may change certain parameters, particularly photolysis rates, to suit their preferences rather than the preferences or recommendations of a mechanism's developer. To examine such differences (and whether they might produce substantive differences in calculated ozone concentrations) would require, among other things, inspection of the actual computer codes, which, even if available, could not have been accomplished in the available time. Third, for the same transport model, results could differ due to the different ways to prepare model-ready emissions for CB-IV and SAPRCC-99, despite starting with the same emissions inventory. Examination of the significance of these differences would have required, at a minimum, examining the emissions processing software for each mechanism.

This review also draws on discussions held on January 10, 2003 from 1:00 – 4:30PM at AQMD headquarters between AQMD and CARB staff as well as other outside reviewers. During this meeting, results of reviewers' initial, preliminary assessments were shared and discussed. It was also during this meeting that reviewers were provided with the results for two additional modeling systems: CALGRID-SAPRC-99 and CAMx-SAPRC-99.

Finally, this review considers the 2003 AQMP modeling from a practical and relatively narrow perspective, at least scientifically. All photochemical grid models, regardless of their science and sophistication, produce estimates of spatially averaged ground-level ozone concentrations for each hour and grid-cell in the region modeled. In the modeling performed for the draft 2003 AQMP, each modeling system produced roughly 200,000 ground-level ozone concentration values for each day modeled – roughly one million values, for example, for the five-day (August 3 – 7, 1997) modeling period.[+] However, in demonstrating attainment of the 1-hour federal ozone standard and in estimating the emissions carrying capacity of the Basin one ozone concentration value has greatest importance and consequence. If this value – the highest of all observed values – is not credibly estimated, then modeling fails. It may seem inefficient to perform all the calculations and work for but a single value, but that is the practice today, and that is the challenge that all modeling systems face in the regulatory world of modeling.

The credibility (or reliability) of model estimates of ground-level ozone concentration comes from comparing calculated and observed values for the historical modeling period (August 3 – 7, 1997). The agreement between observed and calculated values represents something like a figure-of-merit for a modeling system: the better this agreement, the better the modeling system.[++] As noted previously, however, because of the emphasis in decision-making given to the highest ozone concentration, the agreement between maximum calculated and maximum observed values receives extra weight in SIP modeling.

Current practice, expressed somewhat abstractly and in technical terms, is to compare the observed maximum hourly ozone concentration x (which occurred on day D at hour H_x at monitor M) with the predicted maximum ozone concentration y (calculated anywhere in the "vicinity" or subregion of the modeling region that includes M on day D at hour H_y). The statistical expectation is that y should exceed x because the odds are small of the relatively sparse set of monitors recording the actual maximum ozone during any day.[+++] Similarly, y should not be less than x, and if it is, that may be direct evidence that that modeling system is not performing credibly or as well as another modeling system that produces the maximum value x that does exceed y.[◊]

[+] Well over 20 million values are calculated by these models when one accounts for calculations of the above-ground ozone and other pollutant concentrations, which, in complex ways, the model is designed to handle because they affect the ground-level ozone concentrations.

[++] Because model estimates represent spatial average ozone concentrations and the observations are made at specific locations ("points"), one does not expect, a priori, model estimates to agree precisely with the corresponding observations. Moreover, one might expect that such model estimates might, being spatial averages, underestimate the point measurements. However, it turns out that paired observations and predictions agree reasonably well on average, without evidence of bias, provided inputs are of reasonable quality.

[+++] For example, it is worth noting that while the system of monitoring stations in the 2003 AQMP is quite large by any standard, that number is still less than 2 percent of the horizontal grid-cells used in the 2003 AQMP modeling. If one only considers the eastern subregion of the whole modeling region, where the Basin's peak concentrations are observed, the number of monitors in that subregion represent less than 3 percent of that subregion's number of horizontal grid cells.

[◊] Photochemical modeling represents an exercise of a complex modeling system in which the effects of negative biases of one (or more) model component(s) may over- or under-compensate for (or tend to offset) the effects of some positive bias of another model component (or components). The existence of such compensating effects is not only possible but probable although little direct evidence documents the occurrence of its pervasiveness, magnitude or influence in SIP applications. EPA's ozone SIP modeling regulations seek to address this concern through the requirement for "model performance evaluation", an activity in which the

To restate the principal point of the previous paragraph, using less technical terms, a correctly operating modeling system – a credibly operating modeling system – applied for both the purposes of demonstrating attainment of the federal ozone standard in and estimating the emissions carrying capacity of the Basin under meteorological conditions that are conducive to ozone formation, should, among things, calculate maximum hourly ozone concentrations that exceed the corresponding concentrations recorded by correctly operating "nearby" monitoring instruments.

Preliminary Observations, Findings and Recommendations

Observations and Concerns

This new and expanded endeavor to develop the 2003 AQMP for the Basin is a natural and logical extension of AQMD's peerless tradition of comprehensive and effective air quality management planning. AQMD was not only the first to use photochemical grid modeling in the 1980s to develop and implement the nation's first AQMP for ozone, but it has also set the standard for updating those plans. Arguably, AQMD's comprehensive planning is contributing to the Basin's steadily improving ozone air quality, which shows, according to projections of published observations, that the federal 1-hour ozone standard will likely be achieved by 2010, the year specified by law. This record of accomplishment may be taken as evidence of AQMD's ability to effectively account for the multiplying uncertainties associated with air quality management planning.

Notwithstanding AQMD's record of accomplishment, each plan is developed and judged, as noted and described in some detail in the previous section of this review, using the newest available information and latest knowledge. Against this standard, three overarching concerns emerged during the course of this review of the 2003 draft AQMP documentation:

1. The use of multiple modeling systems rather than a "best" modeling system, which deviates from EPA's modeling guidance as well as AQMD's past practices;
2. The apparent use of a relatively small number of episodes all of which represent the same type of high ozone-formation conditions, which also deviates from EPA modeling guidance; and
3. The absence of episodes that include both weekdays and weekends for use in demonstrating attainment and estimating the Basin's carrying capacity, which seems to disregard a well-established Basin trend in

---

veracity of a model system is judged by its ability to replicate important observed features of historical episodes. The regulations presume that if a model's performance evaluation meets EPA requirements, then concerns about any adverse influence from compensating effects are inconsequential.

which peak hourly ozone values – and therefore controlling emissions – occur as or more frequently on weekend days compared to week days.

The importance of these three concerns is the degree to which they undermine the technical credibility of a draft 2003 AQMP, specifically to judgments about the likelihood that the Plan will provide for attainment of the hourly ozone standard by 2010, and overstate or understate, by substantive amounts, estimates of the Basin's carrying capacity.

All three are equally important in the opinion of this reviewer. By not addressing concern one, the Plan faces the criticism of "model shopping", i.e., searching among a set of modeling systems that produces, at least directionally, a predetermined result. Similarly, by not addressing the concerns two and three, the Plan faces the criticism of "episode shopping", i.e., searching for episodes that produce, at least directionally, a predetermined outcome. Additionally, in basing the Plan on but one type of meteorological condition conducive to the formation of high concentrations of ozone, the Plan may not provide control strategies that work for the variety of type of conditions that may be encountered along the way to 2010.

AQMD has sound arguments for not including in its planning the modeling meteorological conditions that rarely occur, such as the July 13 – 18, 1998 period. If, for example, as EPA guidelines permit, episodes have meteorological recurrence times of once in three to five years or less, then those episodes may reasonably be excluded from the demonstration of attainment. The federal 1-hour ozone standard permits the occurrence of concentrations that exceed 0.12 ppm four times in three years. However, after justifiably discarding the rare and "extreme" conditions that are conducive to ozone formation (which the July 1998 period is), the risk remains that, as the spatial, temporal and compositional distributions of emissions change with the implementation and penetration of the Plan's control measures, not modeled meteorological conditions might emerge as controlling, invalidating the Plan.

The third concern – not including weekend periods in the modeling – is, in some ways, a variation of the second concern as well as evidence of its importance. In the case of the third concern, the shift in spatial and temporal emission patterns is already, evidently, leading to recorded shifts in the patterns of high hourly ozone concentrations. By not explicitly modeling these conditions, the Plan is neglecting the risk that the current downward trend in hourly ozone concentrations might begin to slow or even flatten, which could result in the Plan not providing for attainment of the federal hourly ozone standard by 2010. It appears that the only available modeling database that contains a weekend day is also an "extreme" episode, viz., the SCOS97 July 1998 period. Thus, in justifiably discarding the July 1998 episode for its rareness, the only episode with weekend days is also discarded.

However, the modeling for the Plan could be extended in one of two ways to address the risk of weekend events becoming controlling under 2010 emissions. The first and simplest way is to adapt the methods used in the 1997 AQMP. A second way is to use the July 1998 modeling period, not to demonstrate attainment under these conditions, but instead to demonstrate that controlled 2010 emissions based on the August 1997 conditions would provide for attainment under weekend emission conditions.

In this second way, a 2010 controlled emission inventory derived from the same control measures obtained using the August 1997 conditions – except they would reflect weekend activities – would be prepared for the July 1998 episode. The modeling system would be run and the difference noted between a representative weekend day and week day. This difference would be compared with the margin by which the controlled 2010 emissions demonstrated attainment of the standard under the August 1997 conditions.

For example, for a +x ppb difference between the weekend and week days the 2010 controlled inventory would have to be x ppb below 124 ppb, the concentration level of the federal 1-hour standard.

The choice between the two methods is, in the opinion of this reviewer, best left to AQMD. Of the two methods, it could be argued that the latter is better than the former because the July 1998 conditions represent actual weekend periods. However, by the reasoning used to discard the July 1998 conditions for the purposes of demonstrating attainment, one could also argue that the July 1998 conditions are too extreme, and, therefore, likely to overstate the importance of the weekend emissions in judging the risk that the Plan does not provide for attainment by 2010.

If the appearance of model shopping is set aside, there are two additional reasons for concern about model selection or using multiple models. First, EPA's modeling guidance, while allowing for the use of an alternative modeling system, contemplates the use of just one modeling system, not several. That guidance (the still formally applicable 1999 Edition), while providing for EPA approval of an alternative to its "preferred" model, seems also to require that "…the selection … be done in a consistent manner." (40 CFR, Pt. 51, App. W., Sec. 3.2.1a, p. 396) Furthermore, the guidance refers readers to two EPA documents aimed at identifying an approach "…for objective decision-making on the acceptability of [selecting] an alternative model for a given regulatory application." (Ibid.)

Second, EPA has often expressed concern when an alternative model predicts lower episodic peak concentrations than its preferred model (or some other model). EPA's concern includes the situation where a model that excels in its theoretical formulation and that even achieves or exceeds EPA standards of performance underpredicts episodic peak concentrations. (See for example, EPA

Memorandum, “Draft Protocol for the UAM-V—A Correction”, from J. Tikvart to B. Johnson, August 17, 1993) A principal consequence of EPA’s view is that a better model (i.e., an acceptable model) is one that has not only better science and better performance but also does not <u>understate</u> peak episodic concentrations.

However, EPA’s modeling guidance, in describing the use of but one model, does not explicitly exclude, it seems, the possibility of using more than one model. EPA’s formal guidance seems flexible, which squares with its pronouncements at the most recent (seventh) Conference on Air Quality Modeling (June 2000). So, it seems AQMD could use, at least as far as there’s concern for not following EPA guidance, more than one modeling system in its planning. However, throughout its guidance and in the transcripts of its seven modeling conferences, EPA, in considerable detail, expresses its concern for consistency of model use and modeling practice.

If AQMD intends to use multiple modeling systems in its 2003 planning, it follows, then, that it carries a dual burden. First, AQMD should provide arguments that support the use of multiple modeling systems. Second, AQMD should describe the method for using multiple models in planning.

A two-pronged argument might support the use of multiple modeling systems in planning. First, the systems selected should manifestly meet the two EPA tests: (1) meet or exceed EPA model performance goals, and (2) not underestimate observed peak episodic concentrations. Second, the benefits should be sufficient to offset the burden of maintaining and applying multiple models to multiple modeling periods. The use of multiple systems should not be used as justification for modeling fewer episodes. So far, in the material reviewed to date, the argument for using multiple modeling systems has not been presented.

The use of multiple models in planning – the second of the two burdens the AQMD carries in its pursuit of multiple modeling systems – would break new ground in SIP modeling. The immediate and most straightforward approach might be to augment the results obtained using the still formally EPA-preferred system with one or more of the alternative systems. A second more complex approach would be to average the results obtained from those systems that meet EPA’s two tests. However, whether one takes a simple or weighted average requires consideration that is outside the scope of this review.[♦]

The material that AQMD has provided for this review shows that only one model comes close to achieving both EPA tests. As shown in Table 1 (which summarizes the performance of the five systems provided for this review across

[♦] For example, to give more weight to the systems that perform better in estimating observed peak concentrations, an average could be formed by weighting each concentration by its relative ability to reproduce the observations in the subregion where episodic peaks are observed (or some other scheme, which might include combining several metrics of performance).

several performance metrics for the subregion [004] where peak hourly ozone concentrations are routinely observed in the Basin), this model is UAM6.21. The other systems, while evidently meeting EPA's model performance goals, badly understate peak episodic concentrations, by as much as 20 percent in the case of both CAMx-CB-IV for the peak of August 5, 1997 and CALGRID-CB-IV for August 6, 1997. However, CALGRID-SAPRC-99 fairs only slightly better (0.83 vs. 0.80).for the peak of August 6, 1997.• For this reason and by the framework established here, the use of multiple models seems not supportable.

Table 1. Comparisons of Alternative Modeling Systems: Hourly Ozone Concentrations for Subregion 004, August 5 & 6, 1997

Model System	Mean Bias, pphm	Mean Error, pphm	Calculated location of peak relative to observed peak, km	Calculated time of peak relative to observed peak, hour(s)	Correlation with observations	Ratio of calculated peak to observed peak	Calculated peak concentration, pphm
	August 5, 1997						
UAM6.21	-0.6	2.4	33 ESE	3	0.31	0.96	17.9
CALGRID-CB-IV	-0.5	2.5	33 ESE	1	0.45	0.83	15.6
CALGRID-SAPRC-99	1.3	2.7	80 ESE	3	0.46	0.88	16.5
CAMx-CB-IV	0.4	1.8	21 NE	-3	0.57	0.8	14.9
CAMx-SAPRC-99	-0.2	1.9	99 NE	1	0.49	0.86	16.1
	August 6, 1997						
UAM6.21	3.1	1.4	30 NW	1	0.42	1.2	20.4
CALGRID-CB-IV	-0.8	2.7	10 E	-2	0.38	0.8	13.6
CALGRID-SAPRC-99	1	2.8	58 NNW	1	0.43	0.82	14.3
CAMx-CB-IV	-0.1	2.0	46 SW	-3	0.46	0.86	14.9
CAMx-SAPRC-99	-0.3	1.9	70 NE	-2	0.56	0.85	14.5

Source: "Model Performance Evaluation" reports provided by SCAQMD and ARB for each modeling system, with each report based on latest and best available inputs, as of January 16, 2003.

However, should AQMD choose to use multiple systems, then its best present option would appear to be to augment the results of UAM6.21 with the results from either CAMx-CB-IV or CAMx-SAPRC99, or both. These two systems rank high and often using the metrics in Table 1. However, special consideration is warranted because both systems understate peak concentrations by 14 – 20 percent.

• Even if one were to give equal weight to all of the metrics in Table 1 and rank the five systems according to their performance relative the metrics, CAMx-CB-IV would rank first in four of the metrics and UAM6.21 and CAMx-SAPRC-99 would rank first in three, with UAM6.21 best in replicating the peaks.

For example, AQMD might consider following EPA's guidelines that provide for using estimates of the relative change in calculated ozone concentrations rather than the absolute concentrations to evaluate attainment strategies (EPA, 199a, 199b). When using models in this relative way, a system's response to controlled and uncontrolled emissions, expressed as a fraction formed from the ratio of the calculated concentrations relative to the concentrations obtained form the uncontrolled case, is multiplied by the observed concentrations to determine whether the controlled emissions are sufficiently low to provide for attainment. The appropriately formed fraction (called a relative reduction factor, RRF by EPA) is applied to the so-called design value concentration to examine whether the standard is met. Similarly formed fractions may also be applied to locations where other key monitors are located to examine whether a candidate strategy provides for attainment of the standard throughout the Basin.

A note of caution is warranted regarding the expected result from the application RRFs when models understate concentrations. The expected relative response of a system that has already been shown to understate peak ozone concentrations would be to understate the effect that reduced emissions would have on peak ozone concentrations. The practical result, then, of applying EPA's RRFs to such a system would be to lead planners incorrectly toward calling for <u>greater</u> reductions in emissions than would be obtained from a system that did not understate ozone peaks. This may be somewhat counterintuitive, but it follows directly from using a system that does not produce, for some unknown reason(s) at this time, sufficiently high ozone concentrations. If one of two systems cannot produce enough ozone to match the observed ozone in a base simulation, then it should be expected to provide a weaker response to the same emission reduction.

There are of course other methods AQMD might consider for upwardly adjusting results that understate peak concentrations. Applying those presents risks for the AQMP's technical credibility, especially without well developed, technically supportable information about why the systems are understating peak ozone. It is for this and other reasons that EPA has not endorsed using model calibration factors, i.e., factors that one applies to model outputs so that model results "agree" with observations. Similarly, selectively modifying parameters within the system to improve the agreement between selected calculated and observed ozone concentrations without having a supportable technical basis for doing so is not endorsed.

<u>Findings</u>

AQMD seeks, among other things, the answers to four questions, which are provided here.

Question 1: Do the air quality modeling approaches used in the AQMP follow modeling practices conducted by other scientific and regulatory organizations?

In important and practical ways the approaches used in developing the draft 2003 AQMP are ahead of those used by other scientific and regulatory organizations. This reviewer is unaware of any other organization that has used or is attempting (a) to use multiple modeling systems in ozone planning, (b) to evaluate alternative methods for generating model inputs or (c) to compare the results obtained with alternative chemical mechanisms. In pursuing the latest modeling fashions, however, AQMD should be wary of diluting its resources: it should carefully weigh the benefits promised by new "fashions" against the mounting challenges of achieving clean air.

In addition, AQMD should be watchful of loosing its leading national position for identifying effective strategies to achieve healthful ozone levels. Other regions of the country, specifically regions in the southeastern U. S. such as Louisiana, Alabama, Georgia and North Carolina, seem ahead of AQMD in considering strategies for attaining both the 1-hour and the newer federal 8-hour standard. While the 8-hour ozone standard is not a formal consideration of the 2003 AQMP, some consideration should be given to the consonance of measures to attain the 1-hour and 8-hour standards, if for no other reason than to facilitate selection of measures that are similarly effective in lowering 1-hour and 8-hour peak concentrations. There is already evidence that the measures that are lowering the indicators of hourly ozone air quality in the Basin are not as effective in lowering the corresponding indicators of 8-hour ozone air quality. For example, trends in the indicators of 8-hour ozone air quality are dropping at a rate that is approximately two-thirds of the rate of indicators of hourly ozone air quality. Modeling can be a useful tool for identifying those measures that are similarly effective in reducing both 1-hour and 8-hour ozone concentrations.

Question 2: Does the air quality modeling [used in the AQMP] follow regulatory guidelines for application of air quality models?

The answer to this question is mixed. While EPA's guidelines are moving toward providing for flexibility and away from the use of a preferred model, its requirements for a modeling protocol and multi-episode modeling are not changing.

The most recent July 2001 draft of AQMD's and CARB's modeling protocol does not reflect how the final modeling was performed, which is required by EPA guidelines. For example, the July 2001 draft fails to explain, why and when CMAQ was discarded and UAM6.21 added (especially for the post-1987 episode? Why were only two episodes used (or is it only one?) instead of the 4 listed (excluding the July 1998 episode)? What criteria guided the selection of the final meteorological input preparation methods? What version(s) of CB-IV and what fixed-parameter version(s) of SAPRC-99 were used? What were the criteria used for model system selection? In addition, the July 2001 draft of the protocol does not describe the emissions processing procedure for the fixed-version(s) of SAPRC-99 used and how that method compares with the corresponding emissions used in CB-IV. (SAPRC-99 has not been used, to the knowledge of this reviewer, in SIP modeling before.)

The method of episode selection does not follow EPA-recommended procedures. While this reviewer does not dispute the AQMD and CARB's method of selection, the reasons for following a different approach needs mention.

EPA guidelines require using three distinctly different, dominant (non-rare) meteorological types or regimes for modeling. Frequency of occurrence more than the magnitude of the concentration matters. The protocol should address why this requirement is not applicable to the Basin.

When alternative models to EPA' preferred model are contemplated, EPA's modeling guideline (1999 Edition) can be read to require two protocols, which apply (1) to model selection from a set of candidate alternative models, which could be argued to allow for the selection of multiple models (as discussed earlier), and (2) to application of the selected model(s).

Question 3: Is there sufficient information to support the attainment demonstration provided in the Draft AQMP?

With the proviso that the issues raised within this review are addressed and if the draft AQMP follows the structure, scope and format of prior AQMPs, it is reasonable to presume that when the 2003 Draft Plan is prepared, it will support the attainment demonstration.

Consideration should be given to the burden the Plan documentation carries in a time when EPA is proposing flexibility in model selection and model use. It seems that the burden for defending practices – assuring credibility – that have no EPA endorsement and for documenting those practices may be greater when there is no preferred model, i.e., no EPA-approved user's manual or guideline to cite. Having said this, it might also be argued that it might be harder to mount challenges to AQMD-selected practices and to the results obtained therewith when only the "rules" of science apply. Because of the standards of credibility set in previous AQMPs, this reviewer recommends the 2003 Plan aim at more rather than less documentation.

Recommendations

Question 4: What recommendation do you have to improve or enhance future AQMP air quality modeling efforts?

Among scientists that specialize in modeling, including ozone modeling for the purposes of air quality planning, there are many differing opinions about what a best model is (or should be) and how to evaluate models. In answering these questions most modelers would probably consider addressing at least one or more of the following questions: How much weight to place on a model's theoretical formulation? How much weight to place on a model's ability to reproduce observations? How much weight to give to its ease of modification? How much weight to place on the expertise and input data required to apply it?

This reviewer believes there would be a strong consensus on this point: the past provides the best way of validating a model. Unless a model can successfully "predict" the past, it cannot be relied upon for the future. It is from this perspective that the following recommendations are offered.

- Make publicly available criteria that AQMD believes would enhance the technical credibility of its planning. The criteria would seek to steer model system developers toward near- and long-term needs of AQMD for improved technology to model ozone (and PM2.5) and to prepare inputs. The criteria would also identify training needs. The criteria should also send the message that AQMD eschews new technology that adds to the burden of modeling without the demonstrated likelihood of better, more cost-effective planning. AQMD's criteria would emphasize that it seeks technologies that, when compared with the methods it currently uses, can assure less-uncertain results, more quickly obtained and equally uncertain results, or more cost-effective and less or equally uncertain results.

- Subject each new or any modeling system considered by AQMD and its collaborator CARB for use in the Basin, in addition to a comprehensive evaluation of its ability to reproduce all (or a large subset of the) historical episodes available for modeling, to a test of its ability to reproduce the substantial downward trends in observed ozone concentrations that has been observed over the 15-year period 1987 - 2002. This is a direct test of a systems ability to "predict" the past before it is used to predict the future. This has already been successfully done, albeit only partially, for UAM6.21 and would provide strong meaningfully practical information about the veracity and utility of alternative modeling systems. Of course, different systems might require different meteorological inputs (for the same episode) and different methods for processing the same emissions inventory for the same episode. In making this recommendations it is assumed that the best input data are provided for all systems, which is why with this recommendation comes the arguable recommendation that the model system developer "sign-off" on the inputs – in advance of doing the test!.

- Reexamine the scheme for classifying meteorological conditions that are conducive to <u>contemporary</u> ozone formation. The scheme currently used, which was developed in the 1970s and 1980s, may no longer capture the conditions needed for continued progress toward attainment of the 1-hour and 8-hour federal ozone standards. If this reexamination indicates that new meteorological conditions are controlling (or may be controlling), then develop the modeling databases that will support ozone modeling and perform the modeling to identify the measures that effectively lower ozone under the new conditions.

- Understand and resolve issues associated a system's good and marginal ability to reproduce historical episodes before proceeding with its use in planning. The modeling community now offers various methods for accomplishing this, including process analysis, chemical tagging and sensitivity analysis.

- Train staff in the use of the newer methods for generating meteorological inputs. The ingredients of good modeling are technology, inputs and trained staff. While the recipe for good modeling can vary among these ingredients, good meteorological inputs have always been vital to a system's ability to reproduce observed ozone and ozone precursor patterns. With the newer modeling systems, like CAMx, CMAQ and UAM-V, meteorological inputs are obtained using the MM5 meteorological model, which was not originally intended for use with air quality models and which requires considerable experience before it produces reliable and believable results. Among the surprises of this review was the relatively poor ability of the newer models to reproduce peak ozone concentrations, which this reviewer suspects is, without corroborating evidence, attributable in substantial measure to "problems" with the meteorological inputs.

Concluding Comments

The purpose of this review, among others, has been to help assure the path to clean air does not become unproductively cluttered by issues of competing modeling technologies. In the modeling work to support the ozone planning, AQMD and CARB sought, for commendable reasons, to use paving stones formed from and shaped to fit alternatives to the UAM system. In the opinion of this reviewer, the paving stones may fit – according to EPA's modeling performance standards – the alternative systems, but the path formed from these stones then becomes unnecessarily crowded with relatively poorly supported options, which suggest riskier higher carrying capacities and lower emission reductions that could delay clean air.

In the final analysis, however, AQMD has demonstrated, for more than 15 years, its ability to use modeling to navigate a path to clean air. In the early days of planning, the paving stones did not fit UAM very well – to many the view UAM offered seemed surreal, the footing it provided felt uneven, sometimes slippery. It is AQMD's decision to decide whether the fit achieved so far with the alternative systems is good enough. Perhaps the current fit is but an interim fit, adequate for transitioning to alternative systems that will be used in future AQMP updates.

Additional References

EPA 1999a. *Draft Guidance On the Use of Models and Other Analyses In Attainment Demonstrations for the 8-Hour Ozone NAAQS.* EPA-454/R-004, U. S. Environmental Protection Agency, Office of Air Quality Planning and Standards. Research Triangle Park, NC, May 1999.

EPA. 1999b. *Guidance for Improving Weight of Evidence Through Identification of Additional Emission reductions, Not Modeled.* U. S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Emissions, Monitoring, and Analysis Division, Air Quality Modeling Group. Research Triangle Park, NC 27711, November 1999.

Review of Operational Evaluation for South Coast Air Quality Management District 2003 Air Quality Management Plan

Judith C. Chow

January 18, 2003

Introduction

The South Coast Air Quality Management District (SCAQMD) is conducting air quality modeling for ozone (O_3) and PM_{10} in support of its 2003 Air Quality Management Plan. The UAM (Ver 6.21), CALGRID, and CAMx air quality models were applied to 1997 base-year emissions and one field study episode (August 4-6, 1997) to estimate O_3 in southern California for an operational evaluation. The models were configured to be as similar as possible in terms of grid size, boundary conditions, initial conditions, and deposition. The CB-IV chemical mechanism was used in UAM and both the CB-IV and SAPRC99f mechanisms were used in CALGRID and CAMx. The MM5 meteorological model generated prognostic wind fields that drove each of the models. The number and depths of vertical layers differed among the different models.

Operational evaluations compare model outputs to measured concentrations of O_3 and intermediate chemical components. The objective of this operational evaluation is to justify the selection of one of these modeling systems for evaluating effects of planned emission reductions in the South Coast Air Basin (SoCAB). Several comparison statistics and plots were created for O_3, NO_x, and CO model estimates and observations as part of this evaluation.

The review evaluates these performance measures and comments on how these might be used to justify the model selection.

Operational Evaluation Statistics for Ozone

An initial examination of the plots and statistics for the UAM, CALGRID and CAMx models using CB-IV and/or SAPRC99 shows that the SAPRC mechanism yields much better agreement between model-estimated and measured O_3. Given that all other inputs to CALGRID and CAMx are the same, and that the improvement it performance measures is consistent across the two models, it appears that the SAPRC99 mechanism better represents the ozone-forming chemistry in the modeled region, at least on an operational basis. Normalized bias, normalized error, and correlation statistics were tabulated for each sub-region and each of the five model applications. It was clear from this tabulation (not shown here) that biases and errors were smaller and correlations were higher (especially for r>0.6) for most of the sub-regions when the SAPRC99 mechanism was used.

Table 1 summarizes the ranges of performance measures for O_3 classified by the ten sub-regions. Entries in Table 1 are those that were closest to the U.S. EPA (1991) performance objectives of 0.8 t 1.2 for the unpaired peak ratio, ±15% for average normalized bias, and <35% for average normalized gross error. For the models with both mechanisms, the SAPRC99 mechanism better attained these objectives than the CB-IV. However, the UAM with CB-IV was equivalent to or slightly better than the other models for sub-regions 1. N. Central Coast, 3. San Fernando Valley, and 5. LA. It would be useful to determine the extent to which performance measures would improve if the UAM was capable of using both chemical mechanisms as are the CALGRID and CAMx.

All of the models tend to underestimate peak hourly O_3 more often than they overestimate these concentrations. CALGRID-S appears to have the smallest bias most of the time. It usually underestimates O_3 by ~10% to 20% for all sub-regions except at the boundaries, where O_3 is underestimated by ~30%. UAM-C appears to have performed best in the Los Angeles sub-region 5, and comparably with CALGRID-S in the N. Central Coast (2) and San Fernando Valley (3) sub-regions. The CALGRID-S August 6 statistics for the Los Angeles sub-region show a very large bias that contrast with good agreement between calculated and measured values on the previous two days. Valid O_3 measurements were available at only the Pasadena and Palos Verdes monitors in this region, and this may be a partial cause of this and other biases.

The ranges in Table 1 must be qualified by the fact that different numbers of modeled/measured comparisons are included in the comparison statistics. CALGRID-S and CAMx-S comparisons include three more sites (Tijuana, Trona, & Pt. Mugu) than were included for the CALGRID-C and CAMx-C comparison for sub-region 0. It is unclear why the Tijuana site is included in the background category. Pt. Mugu was included in sub-region 2 for the CB-IV comparisons. The Phelan-Beekley site was in sub-region 4 for the CB-IV comparisons and in sub-region 8 for the SAPRC99 comparisons. The sub-region 7 LA N. Main and SJVUCD sites used for SAPRC99 comparisons were omitted for CB-IV comparisons. Documentation of the reasons for including different monitors in each sub-region for CB-IV and SAPRC99 runs is needed. It's not clear why the sub-regions are defined as they are.

Plots of residuals were available only for the CB-IV results. These show that all models tend to overestimate NO at low concentrations and especially at night. They tend to underestimate NO at high concentrations. Residuals appear to be largest for the UAM-C comparisons. NO_2 concentrations biases are smaller than those for NO and shift from overestimation to underestimation at 70 to 80 ppb.

Time series comparisons of calculated and measured O_3 for CALGRID-S and CAMx-S show a tendency to underestimate the peak concentrations. Overestimation is more prevalent at outlying sites (sub-regions 1, 2, 7, 8, and 9). At low concentrations both CALGRID-S and CAMx-S show peaks and valleys that are not similar to those observed in the measurements.

Model Selection for OZONE Control Strategy Simulations

Operational evaluation is only one part of a comprehensive model evaluation effort. Seigneur et al. (2000) recommend: 1) operational testing that demonstrates an ability to estimate O_3 and its intermediate chemical components; 2) diagnostic testing that examines the degree to which precursor and intermediate concentrations are reproduced; 3) mechanistic testing that determines the effects of emission and meteorological changes on estimated concentrations; and 4) probabilistic testing that quantifies uncertainties in model results. The operational comparison statistics indicate that the SAPRC99 mechanism probably performs better than the CB-IV. It is not clear that other differences between the models improve or degrade performance, although the CALGRID-S seems to meet the EPA criteria for most of the sub-regions. In many cases, CAMx-S produced comparable results. The use of more than one model over several episodes would facilitate understanding of the uncertainties related to proposed emission reduction strategies.

PM_{10} Modeling

PM_{10} modeling is not as well documented or evaluated as O_3 modeling. Information given consisted of: 1) tables of a linear rollback calculation, 2) tables of annual average concentrations estimated by UAM-AEROLT and compared with sulfate, nitrate, organic carbon, elemental carbon, ammonium, and “other” concentrations at five sites, and 3) time series plots of calculated and measured concentrations. Ambient measurements are from the 1995 PM_{10} Technical Enhancement Program (PTEP) that partially represent sub-regions 4 and 5 from Table 1.

Comparison statistics show 7% to 34% differences in annual averages between calculated and measured concentrations for the different chemical components. For quarterly averages, the model overestimates measured PM_{10} by 40% to 50% for the first quarter and underestimates PM_{10} by 20% to 50% during the fourth quarter. Part of this might be due to the use of calendar quarters than seasonal quarters (e.g., Dec, Jan, and Feb for winter). Nitrate at Rubidoux is underestimated by 13%, but this includes a 35% overestimate during the first quarter and a 39% underestimate during the fourth quarter. The annual average differences masks more serious differences over shorter time periods. Comparisons of intermediate species concentrations such as ammonia and nitric acid is needed to evaluate the extent to which the model gives the right answers for the right reasons.

There needs to be a greater unity between O_3 and PM_{10} modeling, as NO_x and VOC strategies for O_3 may also affect the ammonium nitrate in PM_{10}. This is another argument in favor of using the SAPRC99 mechanism for ozone, as the CB-IV mechanism cannot be adapted to secondary organic aerosol formation (Stockwell, 2002). It appears that the UAM-AEROLT underestimates secondary organic carbon by factors of 2 to 3 compared to CMB (STMPR presentation, Aug 3, 2001). Several recent studies have used CAMx and CMAQS models for air quality forecasting (e.g., Cai et al., 2002, Stockwell et al., 2002) and for regional modeling in New England and Southern Nevada. Unifying assessment and forecasting capabilities for both O_3 and PM_{10} would be a worthwhile goal that could provide continuous performance evaluation and continued improvement in the modeling/measurement system within the SoCAB.

References

Cai.,C., Hogrefe,C. Biswas, J. Rao,S.T., Seaman, N. Gibbs,A. Kallos, G., Katsfados,P. Walcek, C., Chang J., and Demerjian,K.L. (2002). An experimental air quality forecast modeling system (AWFMS) for the northeast United States: A demonstration study. Presented at the Fourth Conference on Atmospheric Chemisty, Orlando FL, American Meteorological Society.

Seigneur,C., Pun,B.K., Pai,P., Louis,J., Solomon,P., Emery,C., Morris,R., Zahniser,M., Worsnop,D., Koutrakis,P., White,W.H., and Tombach,I.H. (2000). Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility. *JAWMA* **50,** 588-599.

Stockwell, W.R., Artz, RS., Meagher, J.F., Petersen, R.A., Shere,K.L., Grell, G.A., Peckham,S.E., Stein A.F., Piercde, R.V., O'Sullivan, J.M., Whung, P.Y. (2002). The scientific basis of NOAA's air quality forecasting program. *EM* **8**, 20-27.

Stockwell, W.R. (2002). Brief overview of gas-phase atmospheric chemistry mechanisms for air quality models. Internal NOAA report, National Oceanic and Atmospheric Administration, Silver Springs, MD.

U.S. EPA (1991). Guideline for regulatory application of the Urban Airshed Model. EPA-450/4-91-013, U.S. Environmental Protection Agency, Research Triangle Park, NC.

Table 1. Ranges of the best performance statistics for ozone modeling systems by modeling sub-region for the August 4-6, 1997 episode. (More than one modeling system is tabulated when statistics are comparable. Based on reports from Simulation ID arb97b for UAM-C, CALGRID-C, and CAMx-C, p85 for CALGRID-S and p800s for CAMx-S. CALGRID-C and CAMx-C comparisons were examined but they consistently yielded poorer performance statistics than other models.)

Sub-Region[a]	Model[b]	Unpaired Ratio (range)[c]	Percent Normalized Bias (range)[d]	Percent Normalized Gross Error (range)[e]
0. Boundary	CALGRID-S	0.67 to 0.75	-42 to -38	38 to 43
1. N. Central Coast	CALGRID-S	0.89 to 0.95	-14 to 0	8 to 20
	UAM-C	0.8 to 1.02	-19 to -12	16 to 24
2. Ventura County	CALGRID-S	0.82 to 1.24	-4 to 20	11 to 37
	CAMx-S	0.75 to 1.01	-28 to -4	13 to 29
3. San Fernando Valley	CALGRID-S	0.91 to 1.1	-3 to 14	15 to 31
	UAM-C	0.8 to 1.1	-20 to -6	25 to 32
4. Eastern SoCAB	CALGRID-S	0.82 to 0.99	-5 to 16	24 to 34
5. Los Angeles	UAM-C	0.72 to 1.09	3 to 8	22 to 40
Aug 4&5	CALGRID-S	0.81 to 1.0	-10 to -5	12 to 25
Aug 6	CALGRID-S	1.52	-10	60
6. SanDiego/Baja	CAMx-S	0.94 to 0.97	-29 to-26	34 to 38
	CALGRID-S	1.09 to 1.27	-16 to 2	33 to 42
7. S. San Joaquin	CALGRID-	0.82 to	-28 to -14	20 to 42

Valley	S	0.86		
8. Antelope Valley	CAMx-S	0.73 to 1.07	-27 to -8	20 to 28
	CALGRID-S	1.02 to 1.36	3 to 8	22 to 40
9. Imperial Valley	CAMx-S	0.75 to 1.02	-40 to -33	41 to 44
	CALGRID-S	1.13 to 1.60	-3 to 3	26 to 52

[a]Groups of monitoring sites used for model/measurement comparisons.
[b]S designates the SAPRC99 mechanism and C designates the CB-IV mechanism.
[c]Ratio of highest modeled to highest measured O_3 within the sub-region for each modeled day, regardless of hour or location. Range is for the 3-day episodes. Measured O_3 concentrations below 60 ppb are excluded.
[d]Average of differences between modeled and measured O_3 divided by measured O_3 for each day. Measured O_3 concentrations below 60 ppb are excluded. Range is for the 3-day episode.
[e]Average of absolute differences between modeled and measured O3 divided by measured O_3 for each day. Measured O_3 concentrations below 60 ppb are excluded. Range is for the 3-day episode.

21 January 2003

Joe Cassmassi
South Coast Air Quality Management District
21865 Copley Drive
Diamond Bar, CA

Dear Joe,

Here are some preliminary comments and questions for further discussion with respect to the air quality modeling for the South Coast Air Basin based on our meeting on Jan 13. In the short time that was available since that meeting, I was not able to review all of the modeling results or modeling assumptions and input data that were used.

Episode Selection

The air quality modeling effort has focused on the period 3-6 August 1997, which spans a period from Sunday through Wednesday. The highest observed ozone in the SoCAB during this period was 187 ppb at Riverside on 5 August. An advantage to modeling this period is that it occurred during an intensive field sampling phase of the Southern California Ozone Study, so a rich set of air quality and meteorological measurements are available to help define air quality model inputs and to help evaluate model performance.

Key questions:

(1) Are there other more severe air pollution episodes that should be considered when evaluating the emissions-carrying capacity for the air basin? The design value of 187 ppb that was used to determine carrying capacity may be too low. This concern is exacerbated by under-predictions of peak observed ozone levels in the baseline air quality modeling for 3-6 August 1997.

(2) Should an effort be made to consider weekend conditions in the air quality planning effort? For example, NO_x emissions from diesel trucks decrease dramatically on weekends, and ozone increases at some locations in response.

As alternative or additional episodes to be considered, we discussed the period 15-18 July 1998. The highest ozone levels (>240 ppb) were observed at Crestline on July 16 and 18. Other sites such as Glendora, San Bernardino and Redlands also had ozone levels above 210 ppb on July 16, but these sites had peak ozone of 190 ppb or lower on other days. I agree with the staff decision to exclude this extreme event from the design of the air quality management plan.

We also discussed 3 July 1997, which had peak ozone of 205 ppb. You mentioned that there were forest fires during this time and uncertainties in emissions associated with the July 4th holiday. The 3-6 August 1997 episode was preferred because of availability of data from SCOS97 intensive field operations.

We did not discuss the period 29-30 August 1998 which had ozone peaks of 195-200 ppb at Azusa, Riverside, and Glendora. This episode occurred on a weekend.

Another more recent high-ozone event was 25-26 August 2001, with peak ozone levels of 184 ppb at San Bernardino (Aug 25th) and 190 ppb at Glendora on Aug 26th. This episode also occurred on a weekend.

Air Quality Models

I have concerns about the use of the Urban Airshed Model (UAM) and the Carbon Bond IV chemical mechanism in the development of the air quality management plan. In my opinion, better state-of-the-science atmospheric modeling tools are now available. I am not familiar with EPA's policy on what models are acceptable for use in air quality planning. If UAM/CB-IV is used, policy-makers should be informed that this model no longer embodies the "best science".

The meteorological inputs used with UAM-IV differ from those used for the other air quality models (i.e., CALGRID and CAMx). This makes it difficult to compare UAM-IV results to any of the other model simulations. If UAM results are retained, I recommend running a UAM-FCM simulation with the SAPRC99 chemical mechanism, to assess the impact of updating the chemical mechanism without changing other model components. The CALGRID and CAMx simulations seem to give inconsistent results with respect to the effect of switching the chemical mechanism from CB-IV to SAPRC99.

I am uncomfortable with the adjustments made to terrain heights (i.e., reducing all the topographic elevations to 50% of actual height) in the meteorological modeling used to drive the CALGRID and CAMx air quality models. This deliberately introduces a new error to compensate for other problems in the mesoscale meteorological modeling (MM5).

The diagnostic wind model used to interpolate available wind observations to develop wind fields for use with the urban airshed model (UAM) probably provides a reasonable first approximation of flows within the main urbanized area of Los Angeles, but is less suitable for predicting flows through mountain passes for assessment of inter-basin transport.

Carrying Capacity

Peak ozone is under-predicted in all of the air quality model simulations for 3-6 August 1997. In the future year attainment demonstration for 2010, under-prediction of ozone in the base year (1997) is assumed to vanish. I am concerned that 187 ppb may be too low as a design value to start with (see comments on episode selection above), and that the air quality modeling has effectively further "lowered the bar" by starting with ozone levels that are below the observed peak (i.e., 180 ppb for UAM, 154-165 ppb for CALGRID, or 154-161 ppb for CAMx).

In my experience, air quality model results tend to be more reliable when considered on a relative rather than absolute basis. Carrying capacity could be determined based on modeled ozone reductions between 1997 and 2010; these would need to match the change in ozone concentration to go from the design value to the 1-hour ozone standard. For example, if the design value is 190 ppb, and the standard is 120 ppb, the modeling should show a 70 ppb reduction in predicted ozone between 1997 and 2010. Given under-predictions of peak ozone in 1997, a more convincing model-based attainment demonstration would show the needed **change** in predicted ozone between 1997 and 2010.

I plan to provide more detailed comments following further review and discussion of model algorithms, assumptions, and input data. If there are any questions, please call me at (510) 643-9168 or send E-mail to harley@ce.berkeley.edu.

Sincerely,

Robert Harley
Professor

Review of the SCAQMD's 2002 AQMP Ozone Modeling

Fred Lurmann
Sonoma Technology, Inc.
Petaluma, CA 94954
February 7, 2003

1. INTRODUCTION

This review focuses on the ozone air quality modeling that the South Coast Air Quality Management District (SCAQMD) staff performed through December 31, 2002, in support of the 2003 Air Quality Management Plan (AQMP). The scope of the review includes the ozone modeling protocol, the meteorological modeling, and the air quality modeling. The emissions modeling was not reviewed in any detail. The review included examination of the protocol documents, the materials presented to the external modeling review group during 2002, spatial and temporal displays of the meteorological and air quality fields, and tabulated statistics and graphic displays of the model performance. Additional information was provided in a one-day meeting with Joe Cassmassi of SCAQMD, Bruce Jackson of the California Air Resources Board (CARB), and other SCAQMD modeling staff.

2. MODELING PROTOCOL

Draft #7 of the modeling protocol released in July 2001 describes the basic methods used in the modeling. The methods are generally sound and reflect improvements compared to prior SCAQMD modeling protocols. The notable and commendable features of the 2001 protocol are:

- Use of a larger modeling domain than previous studies which minimizes the influence of uncertainties in the boundary concentrations
- Use of state-of-science meteorological models, such as the MM5 prognostic model
- Improved emissions modeling methods and data
- Use of scientifically up-to-date chemical mechanisms and air quality models
- In-depth model evaluation procedures that included multi-species comparisons, subregional comparisons, and aloft comparisons, and that were guided by a philosophy of determining whether the model was getting the right answer for the right reason.

The protocol was overly ambitious with respect to the number of air quality models selected for evaluation. Seven air quality models were selected, including CALGRID, CAMx, CMAQ, SARMAP, UAM-V, UAM/FCM, and UAM-IV. Furthermore, four of the models (Calgrid, CAMx, CMAQ, and UAM\FCM) could be applied using either the Carbon Bond IV (CB4) or SAPRC99 chemical mechanisms. The protocol indicated a preference for using air quality models other than UAM-IV because it was "widely acknowledged to have characteristics which limit its utility when applied to large modeling domains that are not geographically uniform".

The protocol also indicated a preference for using the chemically more up-to-date SAPRC99 chemical mechanism instead of the CB4 mechanism because ARB's Reactivity Scientific Advisory Committee recommended SAPRC99 for SIP modeling. The protocol indicated that the UAM-IV and UAM/FCM would primarily be run for the older (1987 SCAQS) episodes with updated emissions to corroborate findings from the more modern models applied to more recent episodes. The plans and preferences indicated in the protocol are reasonable. In fact, I strongly support the use of more scientifically up-to-date models and mechanisms.

3. EPISODE SELECTION

It is common to select at least three episodes within the last 10 years that had adverse ozone air quality for evaluating the effects of emission control strategies. The SCAQMD selected five episodes for modeling. Two episodes were from the 1987 SCAQS field study (June 24-25 and August 27-28) and were used in the three previous AQMP modeling efforts. Two new episodes (August 3-7 and September 26-29) were from the 1997 SCOS field study. The August 1997 episode was more severe than the September episode, but the September episode included a weekend, which was important for addressing the weekday/weekend ozone differences. The third new episode (July 13-18) was from a period in 1998 with extremely adverse ozone conditions.

The SCAQMD began modeling four of these episodes, but eventually focused on the August 3-7, 1997, episode for the AQMP modeling. I concur that this is the best of the five episodes for evaluating attainment strategies. The July 1998 episode is too stringent and the 1987 episodes are too old to run with modern models. The September 1997 episode received little modeling attention because of difficulties obtaining reliable weekend emissions.

Including a second episode in the evaluation could enhance the robustness of the AQMP modeling. I recommend including an episode with high ozone levels on weekend days like the September 26-29, 1997 period. Even though the range of meteorological conditions causing very high ozone levels in the South Coast Air Basin (SoCAB) has narrowed as emissions have decreased, there is still a need to evaluate control strategy effectiveness on more than one set of adverse conditions.

4. METEOROLOGICAL MODELING

The CALMET diagnostic model was run to provide meteorological fields for the UAM. The MM5 prognostic model was run to provide meteorological fields for CAMx, CALGRID, and CMAQ. The majority of meteorological modeling resources were devoted to the prognostic modeling. In general, the meteorological modeling methods represent improvements compared to earlier AQMP modeling efforts and the new wind fields reflect common patterns for episodes in the SoCAB.

It is difficult (or impossible) to evaluate the meteorological modeling because the details of the methods and model performance are not documented. The SCAQMD/ARB modeling staff should prepare a report documenting the methods, input data, and model performance. It is particularly important to demonstrate an absence of persistent bias in the wind, mixing height, and temperature fields because these will influence the air model simulations. The refinements in the meteorological modeling need to be based on improvements in meteorological model performance, rather than on air quality model performance to avoid potentially compensating biases.

The terrain and the variety of surface types, ranging from ocean to forest and desert, complicate meteorological modeling in Southern California. This type of prognostic modeling is very challenging and there is no recognized "best practice". However, the modeling staff made some choices that raise concerns:

- Four-dimensional data assimilation (FDDA) was applied very aggressively in MM5. Both analysis nudging and observation nudging were applied on all of the spatial scales which can cause excessive forcing of the prognostic fields. Other groups have had good success with a more selective approach to nudging.
- In order to reduce unrealistically high wind speeds in areas of steep terrain with MM5, the terrain heights were reduced by 50%. This is physically unrealistic and inconsistent with modern practice. The wind speed biases were probably symptomatic of other problems that should have been dealt with instead of arbitrarily adjusting the landscape.
- The mixing heights estimated by the MM5 model in the coastal regions were arbitrarily adjusted in post-processing. A key advantage of prognostic models is the physical consistency of all the meteorological fields output by the model. Ad-hoc mixing height adjustments can result in physical inconsistencies, especially with the temperature fields. Adjustments were made to improve agreement with mixing heights diagnosed from vertical sounding. It is likely that improved performance could have been achieved without having to risk physical inconsistency by using one of the alternate planetary boundary layer (PBL) schemes in MM5 or alternate inputs (urban land use characteristics, day-specific ocean surface temperatures, etc.).

Given the rich observational database for these episodes, the CALMET diagnostic model can provide realistic results in areas with relatively smooth terrain. The diagnostic model results in the areas with steep terrain are problematic. In addition, transforming the CALMET model's fixed layer outputs to the variable layers required by the UAM and ensuring mass consistency produces distorted wind fields. As Joe Cassmasssi pointed out in our meeting, transformation of the CALMET fields to the UAM vertical grid causes rotation of the winds, which results in significant errors in the location of the peak concentrations in the air quality model results. Overall, the diagnostic model results have less scientific credibility than the prognostic model results.

5. AIR QUALITY MODELING

The air quality models derive their credibility from their scientific formulation. Specific air quality model simulations derive their credibility not only from the air quality model, but also from the quality of the input data and degree of agreement between predictions and observations. As indicated above, the SCAQMD selected credible models and had relatively good aerometric data to develop model inputs. The models were applied in a reasonably consistent manner. The principal caveat that applies to all of the air quality modeling is that the meteorological models' performance is undocumented and the accuracy of the emission estimates are not well established.

While the SCAQMD/ARB modeling team initially ran numerous models for numerous episodes, the subsequent modeling effort focused on simulations of the August 1997 episode with the CAMx and CALGRID models. Simulations were run using the CB4 and SAPRC99 chemical mechanisms. Simulations were also made with the UAM-IV model to check for consistency with prior modeling efforts, even though the protocol indicated a commitment to use more scientifically up-to-date models for the AQMP modeling. Not surprisingly, the modeling team found significant differences in the ozone and ozone precursor concentrations predicted by different air quality models and chemical mechanisms. They also found differences in the response of the models to emission reductions. While the emissions and boundary conditions were identical for all of the simulations, the differences in the chemical mechanisms, solar radiation and photolytic reaction rates, meteorological fields, deposition algorithms, vertical grid structure, and numerical algorithms combined to produce different results.

Simulations with the more chemically up-to-date SAPRC99 chemical mechanism produced slightly higher ozone concentrations in the CAMx and CALGRID models than simulations with the CB4 mechanism. The CALGRID model simulations generally produced higher ozone concentrations than CAMx model simulations. Both models underestimated the maximum 1-hr ozone level on the peak day (161 ppb in CAMX/SAPRC99 or 165 ppb in CALGRID/SAPRC99 vs. 187 ppb measured). While the peak ratio statistic is important for regulatory purposes, it is less statistically robust than the mean normalized bias (MNB) and mean normalized error (MNE), which measure performance throughout the modeling domain. The CAMx/SAPRC99 simulations have less bias and error in ozone predictions, on average, than the CALGRID/SAPRC99 simulations. Both simulations meet the U.S. Environmental Protection Agency's (EPA) statistical performance guidelines in the high ozone subregion. I believe both simulations are equally credible.

The UAM-IV simulations predict significantly higher peak ozone levels than comparable CAMx/CB4 and CALGRID/CB4 simulations. The peak ozone is very sensitive to the solar radiation and photolytic reaction rates. The UAM-IV simulations were made with unjustifiably high (+15%) photolytic reaction rates and, therefore, are not directly comparable. If the SCAQMD wants to include the UAM-IV model in the evaluation, then new simulations with comparable photolysis rates need to be run. In fact, for future modeling studies I recommend protocols that stipulate the use of consistent solar radiation and photolytic rate inputs to all of the models because this is a very important forcing factor.

The SCAQMD staff expressed concern regarding the underestimation of the basinwide peak concentrations in all of the models and how this should influence model acceptance. I think the basinwide 1-hr maximum is like an extreme value statistic. It is very difficult to predict it accurately, especially in space and time. Uncertainties in the models, emissions, meteorology, and chemistry contribute to this difficulty. A failure to capture unusual events in the emission inventory could contribute to this problem. Also, the general tendency of having too much NO_x and not enough volatile organic compounds (VOCs) in the simulations may reflect bias in the emission inventory that could clearly affect peak ozone performance more than average performance. As indicated above, I do not think performance on the peak concentration is as important in acceptance of simulations as the MNB and MNE statistics and the reasonableness of the spatial patterns of ozone and ozone precursors.

The models estimate different emissions carrying capacities for the National Ambient Air Quality Standards (NAAQS) attainment scenario. The carrying capacity is defined here as the amount of anthropogenic VOC emissions that would result in attainment of the 1-hr NAAQS (or 124 ppb) when NO_x emissions were reduced to 560 tons/per day (tons). Based on “across-the-board” emission reductions, the CALGRID/SAPRC99 and CAMx/SAPRC99 simulations estimate 340 and 500 tons carrying capacities, respectively. Carrying capacity estimates based on actual emission control measures plus a “black box” of unknown control measures indicate slightly lower amounts (e.g., 300 vs. 340 tons). The carrying capacity differences are consistent with the model performance. The carrying capacity calculation uses the model results in an absolute manner. CAMx/SAPRC99 estimates lower ozone levels in the baseline simulation (on average and at the peak station) and, therefore, requires smaller VOC reductions to achieve attainment than comparable simulations with the CALGRID/SAPRC99. Both estimates are credible; they indicate that the carrying capacity based on “across the board” reductions is 420 ±80 tons. Agreement of ±20% is not unreasonable considering the modeling results are being interpreted in an absolute sense.

I recommend that the SCAQMD staff follow their protocol and also calculate the Relative Response Factor (RRF) for these simulations. The model results are used in a relative sense with the RRF approach. I suspect both models will estimate lower carrying capacity with the RRF approach, but will be in closer agreement. Both the absolute and relative approaches have technical merit. One hopes to find convergences of the results to within ±10 or 15%.

6. THE 8-HR NAAQS

There is conspicuous omission of 8-hr ozone standard in this process. The modeling effort seems to ignore the need to examine compliance with the 8-hr NAAQS in the next few years. The episode selection, model performance evaluation, and control strategy evaluation are strictly focused on the 1-hr standard. The SCAQMD should consider more integrated planning approaches that strive to develop plans to meet both ozone standards. At a minimum, the daily 8-hr maximum performance statistics and effectiveness of emission control measures in achieving compliance with the 8-hr standard should be documented. This information will be informative even though the August 1997 episode may not be the best episode for evaluating control strategies to meet the 8-hr NAAQS. This is a good opportunity to educate the public about the challenges of meeting the 8-hr standard and also show them approximately how much progress towards compliance with the 8-hr standard can be achieved with the current plan.

Review, Critique and Recommendations Concerning Modeling of Ozone and Particulate Matter in the South Coast Air Quality Management District

By

Edwin L. Meyer, Jr., PhD, PE

Prepared in partial fulfillment of AQMD Contract Number 03223

January 27, 2003

Acknowledgment of Support and Disclaimer

This report was prepared as a result of work sponsored, paid for, in whole or in part, by the South Coast Air Quality Management District (AQMD). The opinions, findings, conclusions, and recommendations are those of the author and do not necessarily represent the views of AQMD. AQMD, its officers, employees, contractors, and subcontractors make no warranty, expressed or implied, and assume no legal liability for the information in this report. AQMD has not approved or disapproved this report, nor has AQMD passed upon the accuracy or adequacy of the information contained herein.

Table of Contents

1.0 Introduction

This review is based on materials provided by the South Coast Air Quality Management District (AQMD) as well as discussions with the AQMD and other reviewers in a meeting on January 10, 2003. Prior to the January 10 meeting, I was provided with the following information:

1. an initial modeling protocol;
2. peer review comments on the initial protocol;
3. a series of presentations by the technical staff of AQMD and/or California Air Resources Board (CARB) to stakeholders;
4. a draft appendix describing current and future inventories, as well as methods used to make these estimates;
5. a brief discussion addressing model selection;
6. results of performance tests for ozone and precursors using CALGRID, CAMx and UAM6.21, all with the carbon bond 4 (CB4) chemical mechanism;
7. results of performance tests performed for 6 components of particles less than or equal to 10 micrometers aerodynamic diameter (PM10) and particles less than or equal to 2.5 micrometers aerodynamic diameter (PM2.5) using the UAM-LT air quality model, and
8. projected estimates in 2006 and 2010 for PM10 and PM2.5 and their components.

Subsequent to the January 10 meeting, I was also provided with the results of performance tests for ozone, N0x and CO obtained using the CALGRID and CAMx models with the SAPRC99f chemical mechanism (rather than the CB4 mechanism).

Much of the AQMD's concern is focused on whether the modeling approaches examined are adequate to support the 2003 State implementation plan (SIP) revision. In particular, they wish to know which, if any, of the 5 modeling approaches considered for ozone (UAM/CB4, CALGRID/CB4, CAMx/CB4, CALGRID/SAPRC99f, CAMx/SAPRC99f) is likely to be the most reliable approach. The AQMD has asked reviewers to respond to 8 questions regarding each of the 5 ozone modeling approaches under consideration. AQMD also poses three additional, general modeling questions for us to respond to. In Section 2.0, I describe results of my review of ozone model performance and make a recommendation regarding choice of a modeling approach for the 2003 SIP revision. Section 3.0 addresses questions posed by the AQMD. Section 4.0 discusses information presented about PM10 and PM2.5. Section 5.0 makes recommendations concerning future analyses and model performance evaluations. Section 6.0 summarizes material presented in the preceding sections.

2.0 Performance Evaluation for Ozone and Precursors

Model performance information I review covers a 3-day period (August 4-6, 1997). This period corresponds to an intensive monitoring study (SCOS97) and is one of the episodes modeled for the 2003 SIP revision. The performance evaluations for ozone include several aggregate statistical measures: ratios of unpaired global predicted to observed peak concentrations, ratios of spatially paired peaks, normalized bias, gross error and time series correlations between observations and predictions. All such comparisons are compartmentalized into 9 zones within a large, regional modeling domain. Demarcation of the zones is somewhat subjective. It is based on location of primary sources of precursors, as well as locations of high observed ozone. Other considerations, like assurance of a sufficient number of air quality monitors in each zone and some correspondence with neighboring air quality basins, which in the past have been modeled separately, also appear to apply.

Graphical displays constitute the other major means for evaluating model performance. The displays include diurnal comparisons of predicted and observed ozone concentrations at monitoring sites. Diurnal curves are available for ozone for all five modeling approaches. For the 3 approaches featuring the CB4 mechanism, diurnal curves are also available for NO, NO2 and CO. Stratified curves, examining dependencies between model performance and level of observations are also available.
For the two approaches featuring the SAPRC99f mechanism, all that is available are diurnal curves for N0x, CO and ozone.

My review proceeds by first considering how modeling results have usually been interpreted by the U.S. EPA to determine whether attainment of the national ambient air quality standard (NAAQS) for ozone will be achieved by a required date. Based on this, I emphasize several performance measures for ozone, which, in my opinion, appear closely related to the predictions likely to receive greatest attention in the attainment demonstration. I then combine these performance measures in such a way to develop a numerical rating for each of the 5 modeling approaches. This procedure produces an initial choice. Next, I compare the initially chosen model's ability to replicate observed precursor concentrations against that for alternative approaches. Markedly worse performance would be cause for reexamining the initial choice.

2.1 Model Performance Predicting Ozone

I used six criteria to evaluate performance and weighted each as shown.

1. Normalized bias (all sites in each zone) $\leq \pm 15\%$. (BIAS) (1 point)

2. Gross error (all sites in each zone) $\leq 30\%$. (ERROR) (1 point)

3. Most sites, including the site with the highest observed peak have predicted and observed peaks within $\pm$ 20%. (SITE PEAK) (1 point)

4. Within a zone, the predicted global peak is at least as high as the observed global peak, but not more than 20% higher than the observed global peak. (GLOBAL PEAK) (2 points)

5. Within a zone, the time of the predicted global peak is within 2 hours of the observed global peak and the same is true for most site-specific peaks. (TIMING) (1 point)

6. Within a zone, predicted exceedances of 124 ppb coincide with observed exceedances (EXCEEDANCES)
   (a) at all sites with an observed or predicted exceedance (2 points)
   or
   (b) at most sites with an observed or predicted exceedance. (1 point)

For any given zone, a model could have a score as high as 8 points on each of the modeled days, or 24 points for the 3-day episode. In a zone where there are no observed or predicted exceedances at any site, the maximum daily score is 6 points, or 18 points for the episode.

Criteria 1 and 2 are similar to ones suggested by the U.S. EPA (U.S.EPA, 1991, 1996). They address, all predicted/observed pairs, not just cases where predictions and/or observations are high. This provides some assurance that the model is working well under variety of conditions. The second criterion puts a limit on the size of residual error between predictions and observations, while the first criterion ensures that there is a balance between over- and underpredictions.

Criterion 3 emphasizes performance in predicting peak hourly ozone concentration at each site with observations. Peak predicted ozone usually receives great emphasis in attainment demonstrations.

The fourth criterion also emphasizes ability to predict peak hourly ozone concentrations accurately. Predicted global maximum is often the key factor in an attainment demonstration. For attainment to be shown, it needs to be $\leq$ 124 ppb or to be explained and discounted through a weight of evidence determination. Because there are many more surface grid cells than monitoring sites, the criterion requires predicted global maxima to be at least as high as observed global maxima. Since the South Coast Air Basin has a dense array of monitors, an upper bound has been added for the predicted global maximum ozone concentration to suggest adequate performance. Because the predicted global maximum is such an important factor in an attainment demonstration, satisfactory performance according to this criterion is awarded 2 points.

Criterion 5 addresses the timing of predicted vs. observed peak ozone concentrations. Poor timing could be due either to inadequate meteorological characterization or to some flaw in the modeled chemistry. If a chemical problem exists, large delays in predicted peaks could overstate the benefits of VOC control. In contrast, if predicted peaks occur well before the observed ones, benefits of N0x control might be overstated. If the poor timing is attributable to meteorological problems, incorrect combinations of sources could be mixing at inappropriate times, leading to a potential error in the predicted effectiveness of controls.

Criterion 6 addresses how well each of the models performs in predicting concentrations that approach the level of the NAAQS. This is likely to be a critical issue in the "post-control" state, with relatively small differences resulting in potentially large differences in the cost of control efforts. Thus, consistently accurate performance in meeting this criterion is given 2 points. If the model gets this measure right most of the time, 1 point is awarded. Note that a very small difference (e.g., 125 ppb predicted vs. 124 ppb observed) precludes a model from getting the full 2 points.

Tables 2.1 – 2.5 show performance results for each of the five modeling approaches. I summarize the results from Tables 2.1 – 2.5 in two ways. The first considers results from all 9 zones. However, in several of the zones monitoring sites are spread out over large areas. Further, these zones may experience lower ozone concentrations and are probably not as critical for determining whether the South Coast Air Quality Management District will attain the NAAQS. Thus, model performance is also summarized for zones 2-5. The summary is shown in Table 2-6. Numbers in the body of the table depict the average daily score per zone for each group of zones.

Table 2.6. Summary of Model Performance Predicting Ozone

Model	Zones 1 – 9	Zones 2 – 5
UAM/CB4	2.36	2.73
CALGRID/CB4	2.40	3.45
CAMx/CB4	2.44	3.09
CALGRID/SAPRC99f	3.24	4.45
CAMx/SAPRC99f	2.76	3.82

It is apparent from Table 2.6 that there is little to choose among the modeling approaches when the CB4 mechanism is used and all 9 zones are considered. Focusing only on zones 2-5 with CB4, CALGRID begins to emerge as having the best performance. Substituting SAPRC99f for CB4, substantially improves the performance of CALGRID. Performance of CAMx is also considerably improved. Although a version of UAM exists which permits substituting a different chemical mechanism, this version was not used by AQMD. Thus, it is not possible to note the effects using SAPRC99f has on the UAM's predictions. Table 2.6 indicates that CALGRID/SAPRC99f is the best performing model for predicting observed ozone on the three days tested.

Table 2.1. Performance of UAM/CB4 Predicting Observed Ozone

Zone	August 4, 1997	August 5, 1997	August 6, 1997
1	Sample too small	Bias-1 pt. Error-1 pt. Site peak-1 pt.	Bias-1 pt. Error-1 pt. Site peak-1 pt.
2	Error-1 pt. Site peak-1 pt. Global peak-2 pts.	Error-1 pt. Global peak-2 pts.	Bias-1 pt. Site peak-1 pt. Exceedances-1 pt.
3	Bias-1 pt. Error-1 pt. Timing-1 pt.	0 pts.	Bias-1 pt. Site peak-1 pt.
4	Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.	Bias-1 pt. Error-1 pt. Timing-1 pt.	Bias-1 pt. Global peak-2 pts. Exceedances-1 pt.
5	Bias-1 pt. Error-1 pt. Timing-1 pt.	Site peak-1 pt.	Sample too small
6	Site peak-1 pt. Timing-1 pt.	Bias-1 pt. Error-1 pt.	Bias-1 pt. Timing-1 pt.
7	Error-1 pt. Global peak-2 pts.	Site peak-1 pt.	0 pts.
8	Error-1 pt. Timing-1 pt.	Error-1 pt. Global peak-2 pts.	Bias-1 pt. Error-1 pt. Timing-1 pt. Exceedances-2 pts.
9	Bias-1 pt.	Bias-1 pt. Site peak-1 pt.	0 pts.

Total = 59 points.

Table 2.2. Performance of CALGRID/CB4 Predicting Ozone

Zone	August 4, 1997	August 5, 1997	August 6, 1997
1	Sample too small	Error-1 pt. Timing-1 pt.	0 pts.
2	Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt.	Timing-1 pt.	Bias-1 pt.
3	Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt.	Bias-1 pt. Error-1 pt. Timing-1 pt.	Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.
4	Bias-1 pt. Error-1 pt. Site peak-1 pt Global peak-2 pts. Timing-1 pt.	Bias-1 pt. Error-1 pt. Timing-1 pt.	Bias-1 pt. Timing-1 pt.
5	Bias-1 pt. Error-1 pt. Timing-1 pt. Global peak-2 pts.	Site peak-1 pt. Timing-1 pt. Global peak-2 pts.	Sample too small
6	Bias-1 pt. Timing-1 pt.	Timing-1 pt. Site peak-1 pt.	Site peak-1 pt. Global peak-2 pts.
7	Error-1 pt. Timing-1 pt.	Timing-1 pt.	Timing-1 pt.
8	Bias-1 pt. Timing-1 pt.	Global peak-2 pts.	Timing-1 pt. Exceedances-1 pt.
9	0 pts.	0 pts.	Site peak-1 pt.

Total = 60 pts.

Table 2.3. Performance of CAMx/CB4 Predicting Ozone

Zone	August 4, 1997	August 5, 1997	August 6, 1997
1	Sample too small	Timing-1 pt.	0 pts.
2	Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt.	0 pts.	Bias-1 pt. Site peak-1 pt.
3	Bias-1 pt. Error-1 pt. Timing-1 pt.	Bias-1 pt. Error-1 pt. Timing-1 pt.	Bias-1 pt. Error-1 pt. Site peak-1 pt. Exceedances-2 pts.
4	Timing-1 pt.	Bias-1 pt. Error-1 pt. Timing-1 pt. Exceedances-1 pt.	Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.
5	Bias-1 pt. Timing-1 pt.	Bias-1 pt. Error-1 pt. Timing-1 pt. Global peak-2 pts.	Sample too small
6	Global peak-2 pts.	Error-1 pt. Global peak-2 pts.	Error-1 pt. Site peak-1 pt. Timing-1 pt.
7	Timing-1 pt.	Timing-1 pt.	Timing- 1 pt.
8	Bias-1 pt. Error-1 pt. Timing-1 pt.	Global peak-2 pts.	Bias-1 pt. Error-1 pt. Global peak-1 pt. Timing-1 pt.
9	0 pts.	Global peak-2 pts.	Global peak-2 pts. Timing-1 pt.

Total = 61 pts.

Table 2.4. Performance of CALGRID/SAPRAC99f Predicting Ozone

Zone	**August 4, 1997**	**August 5, 1997**	**August 6, 1997**
1	**Sample too small**	**Bias-1 pt. Error-1 pt. Timing-1 pt.**	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.**
2	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak-2 pts.**	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt.**	**Timing-1 pt. Exceedances-1 pt.**
3	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak- 2 pts. Timing-1 pt.**	**Bias-1 pt. Site peak-1 pt. Global peak-2 pts. Exceedances-1 pt.**	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt. Exceedances-1 pt.**
4	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt. Exceedances-1 pt.**	**Site peak-1 pt. Exceedances-1 pt.**	**Bias-1 pt. Site peak-1 pt. Timing-1 pt. Exceedances-1 pt.**
5	**Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.**	**Bias-1 pt. Error-1 pt. Site peak-1 pt.**	**Sample too small**
6	**Bias-1 pt. Site peak-1 pt. Timing-1 pt.**	**Bias-1 pt.**	**Bias-1 pt. Site peak-1 pt.**
7	**Bias-1 pt. Error-1 pt. Site peak-1 pt.**	**Error-1 pt. Site peak-1 pt. Timing-1 pt.**	**Error-1 pt. Timing-1 pt.**
8	**Bias-1 pt. Error-1 pt. Timing-1 pt.**	**Bias-1 pt.**	**Bias-1 pt. Error-1 pt. Timing-1 pt.**
9	**Bias-1 pt.**	**Bias-1 pt. Error-2 pts.**	**Bias-1 pt.**

Total = 81 pts.

Table 2.5. Performance of CAMx/SAPRC99f Predicting Ozone

Zone	August 4, 1997	August 5, 1997	August 6, 1997
1	Sample too small	Timing-1 pt.	Error-1 pt.
2	Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.	Error-1 pt. Timing-1-pt.	Bias-1 pt. Error-1 pt. Site peak-1 pt. Exceedances-1 pt.
3	Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt.	0 pts.	Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt. Exceedances-1 pt.
4	Bias-1 pt. Error-1 pt. Site peak-1 pt. Global peak-2 pts. Timing-1 pt. Exceedances-1 pt.	Bias-1 pt. Error-1 pt. Exceedances-1 pt.	Bias-1 pt. Error-1 pt. Site peak-1 pt.
5	Error-1 pt. Global peak-2 pts. Timing-1 pt.	Error-1 pt. Global peak-2 pts. Timing-1 pt.	Sample too small
6	Site peak-1 pt. Global peak-2 pts.	Global peak-2 pts. Timing-1 pt.	Site peak-1 pt. Global peak-2 pts. Timing-1 pt.
7	Timing-1 pt.	Timing-1 pt.	Timing-1 pt.
8	Bias-1 pt. Error-1 pt. Site peak-1 pt. Timing-1 pt.	Error-1 pt.	Error-1 pt. Global peak-2 pts. Timing-1 pt.
9	0 pts.	Global peak-2 pts.	Site peak-1 pt.

Total = 69 pts.

2.2 Model Performance Predicting Precursors and Other Considerations

Agreement between predicted and observed primary precursors (NO and CO) is poor for all three models which use the CB4 mechanism. Only N0x prediction/observation comparisons are provided for CALGRID/SAPRC99f and CAMx/SAPRC99f. However, observed data, as well as data with the 3 models using CB4, indicate NO peaks occur in the early morning, usually about 6 – 9 am local time. NO2 peaks, when they occur, usually happen about 11 am – noon local time. Thus, it is possible to infer whether it is NO or NO2 which influences predicted N0x concentrations at various times of day. Generally, like the other 3 models, those which use SAPRC99f do not reproduce observed concentrations of NO very well.

Based on composite sets of predicted/observed NO comparisons presented for each of the three models using CB4, the CALGRID and CAMx models tend to underpredict highest observed NO and overpredict the lowest observed concentrations. This tendency is a little more pronounced for the CALGRID model than it is for CAMx. There is much scatter between observations and predictions within the range of observations. In my opinion, it is likely that the poor performance of the CALGRID and CAMx models predicting NO (and CO) is attributable to the incommensurability problem. That is, sources of NO and CO are not uniformly distributed, and monitored data are likely to vary more than predictions, because they are subject to small scale meteorological and emissions variations which are not reflected by the uniform model inputs and outputs spread over 5 km grid cells of various depths.

In contrast to the CALGRID/CB4 and CAMx/CB4 models, UAM/CB4 almost uniformly overpredicts observed NO, sometimes by considerable amounts. Since NO peaks tend to be observed in the early morning (before much chemistry has happened), I believe the contrast among the three models is most likely a reflection of the first few vertical layers assumed in the three approaches. Perhaps in the early morning, the first layer is deeper in the UAM approach, allowing for more uniform mixing of NO, which is seldom reflected by the monitored data.

Performance improves for all three models when paired comparisons of predicted and observed concentrations of NO2 are considered. There is a tendency for each of the models to underestimate highest observed concentrations of NO2 (say above 5 pphm). This tendency is a little more pronounced for CALGRID/CB4 than it is with the other two models. Performance of CAMx/CB4 and UAM/CB4 in predicting NO2 appears about the same.

All we are able to consider with CAMx/SAPRC99f and CALGRID/SAPRC99f are N0x, rather than NO and NO2 data. I examine zone 5 (a source-intensive area) and zone 4 (an area where highest peak ozone concentrations are often observed). Numerical performance statistics for the ability of these two models to predict observed N0x in the two zones are not particularly meaningful, since the predictions and observations reflect a mixture of NO and NO2. Timing of predicted peaks (probably largely attributable to NO) is better for CAMx/SAPRC99f than for CALGRID/SAPRC99f. Although the time series correlation coefficients between observations and predictions are not especially good, they are systematically higher for CAMx/SAPRC99f than for CALGRID/SAPRC99f.

Because the numerical comparisons between observed and predicted N0x are misleading (due to N0x being a mixture of NO and NO2), I rely primarily on graphical data to assess performance of the models in replicating observed concentrations of NO and NO2. Thus, my assessment of performance is more subjective than it is for ozone.

I compare performance of CALGRID/SAPRC99f with that of CAMx/SAPRC99f and UAM/CB4 at 8 sites in zone 5: Anaheim, Hawthorne, Los Angeles, La Habra, Lynwood, Long Beach, Pasadena and Pico Rivera. These sites are in source intensive areas, and monitored observations may be higher (and perhaps more reliable) than elsewhere. I consider factors such as ability of predicted peaks to track observed patterns at each of these sites, as well as the extent to which levels of observed peak N0x concentrations are captured. In the case of the UAM/CB4, I have to "eyeball" a sum of NO and NO2 observations. Often, the differences between models are either small, or there are compensating features (e.g., one model predicts diurnal patterns better, while another captures an observed peak concentration more accurately).

Table 2.7. Selected Comparisons of N0x Performance

Location	CALGRID/SAPRC99f vs. UAM/CB4	CALGRID/SAPRC99f vs. CAMx/SAPRC99f
Anaheim	CALGRID	CAMx
Hawthorne	CALGRID	Even
Los Angeles	CALGRID	CAMx
La Habra	Even	CALGRID
Lynwood	UAM	Even
Long Beach	Even	CAMx
Pasadena	CALGRID	Even
Pico Rivera	Even	CAMx

One might infer from Table 2.7 that adding the SAPRC99f mechanism to CALGRID improved its performance vis a vis UAM. However, the conclusions in the table may simply reflect the finding that UAM appears to do a much worse job predicting NO and a slightly better job predicting NO2. Results of the comparison between CALGRID/SAPRC99f and CAMx/SAPRC99f are consistent with the earlier finding that CAMx/CB4 does slightly better than CALGRID/CB4 replicating observed concentrations of NO and NO2. It appears as though predicted N0x disappears more slowly with CAMx/SAPRC99f than it does with CALGRID/SAPRC99f.

2.3 Recommendations

CALGRID/SAPRC99f does the best job predicting observed ozone. On the other hand, CAMx/SAPRC99f does a somewhat better job predicting observed NO and NO2. Both models that use SAPRC99f do a better job predicting ozone than any of the three models using CB4.

I recommend that AQMD use CALGRID/SAPRC99f as its primary model to support the 2003 SIP revision. This recommendation is based on the following:

- the difference in reproducing key ozone observations is more pronounced than the differences in reproducing observed NO and NO2;

- I have more confidence in the representativeness and validity of the observed ozone data than I do in the NO and NO2 data, and the method I use to compare performance vis a vis NO, NO2 and N0x is more subjective than that used to assess performance replicating observed ozone;

- past peer reviewers have strongly recommended using the SAPRC99f mechanism over CB4 on the basis of scientific merit, and it appears feasible for the AQMD to do so;

- staff of the California Air Resources Board has had considerably more experience using CALGRID than CAMx;

- I think it will be essential to use a diagnostic wind model (CALMET) to augment coarse scale MM5 meteorological inputs. I have concerns about the feasibility of using fine scale MM5 predictions alone, or even with FDDA in a complex area like the AQMD. It is probably easier to interface CALMET with CALGRID than it is with CAMx.

3.0 Questions Posed by the South Coast Air Quality Management District

The AQMD has asked each reviewer to address 8 questions to each of the five models that they are considering for use.

1. Is the modeling protocol adequate for the proposed attainment demonstration? Does it require revisions to satisfy EPA modeling guidance?

Most current (draft) EPA guidance (U.S. EPA, 1999) for ozone modeling identifies 15 topics that should be addressed in the protocol. Many of these, are in fact, addressed. Those which are not include (1) identifying the stakeholders involved with reviewing results and the procedure followed in revising the initial protocol as the analysis proceeds; (2) types of analyses used in weight of evidence (WOE) determinations, if WOE is used; (3) procedures used to archive, document and report results; (4) identification of specific deliverables and schedule for delivery to U.S. EPA Region IX.

The preceding omissions notwithstanding, I think the AQMD developed a good initial protocol, which provided an adequate basis to proceed with the analysis. Further, evidence is presented to show that there were numerous briefings in which subsequent results were discussed and proposed changes to initial ideas were pursued. The main concern I have about the protocol is that it is no longer clear what was *actually done* to produce the results whose performance we reviewed. The procedures that were finally followed need to be documented. It would also be helpful to mention how these differ from what was originally proposed. For example, I was uncertain about what was finally assumed regarding vertical structure of the atmosphere in CALGRID vs. CAMx vs. UAM. This makes it difficult to come up with explanations about why predicted and observed precursor concentrations sometimes disagree.

2. Is the science embodied in the model and mechanism adequate for use in the AQMP? If not, why not?

The science in the UAM is older than that in CALGRID and, especially, CAMx. However, if UAM/CB4 had outperformed the other 4 approaches or it had not been feasible to implement any of the newer approaches, I would not have hesitated to recommend its use. After all, it has been used in many other applications. In the same vein, the CB4 mechanism is older than SAPRC99f, and I believe a consensus of prominent scientists has stated that SAPRC99f is the most current chemical mechanism available. CB4 has advantages in that it is less resource intensive to use and is consistent with many existing emissions databases. I would not have hesitated to recommend it if practical considerations had precluded use of SAPRC99f in concert with a grid model. All of the models being considered have to make a series of assumptions. It is the *combination* of these assumptions which produces predicted air quality values. Although some features of a particular model may reflect a more current scientific understanding, it does not necessarily follow that all of its assumptions are correct. Indeed, it is possible for performance to deteriorate if a less appropriate *combination* of assumptions results.

3. Are the meteorological models, pre- and post processors adequate for using this model/mechanism in the AQMP?

At the January 10 meeting, we were assured by CARB and AQMD technical staffs that it was feasible to use any of the 5 models within the timeframe required for preparing the 2003 SIP revision. Although I am not sure what post-processors the question is alluding to, I assume that since any of the 5 approaches can be used to produce results within the required time frame, they are adequate. One reservation I have is with the time and resources to run MM5 on a fine (<~8-10 km) scale with many vertical layers, as well as how realistic such predictions are likely to be in a complex terrain area like the south coast. This is one of the factors I considered in making the recommendation I did in Section 2.3.

4. Was the model applied (e.g., initial and boundary conditions, modeling domain, number layers, etc.) in a manner that would make it adequate for use in the AQMP?

I think the initial protocol and subsequent technical presentations gave evidence that these issues were addressed thoughtfully. The initial assumptions, as described in the protocol, seem plausible to me. I am reasonably confident that these issues were considered in a technically competent manner and that changes which were made were warranted technically. As I noted in response to question 1 however, it is not clear *how* the models were finally applied, given the various changes that became adviseable in the evolving analyses.

5. Is the sub-regional zone definition used to group model performance adequate to characterize model performance for ozone and precursors?

Dividing the domain into zones is a very good idea and is, incidentally, consistent with a recommendation in the most current U.S. EPA ozone modeling guidance (U.S. EPA, 1999). It has not been done too often, because most locations do not have a sufficient number of monitors to permit robust comparisons. With its large network, complex topography and geographic distribution of precursor emissions and observed high ozone, the AQMD is well advised to divide the area into zones. The zones are appropriate in that they appear to be based on past ozone observations, distribution of emissions, topography and distinctive air basins which heretofore have been considered separately.

6. Are the graphical ozone model performance presentations consistent with the statistical evaluation?

I don't believe I could have done a meaningful evaluation without the statistical *and* the graphical presentations. The graphical presentations are helpful for trying to understand what is happening to cause the numbers generated by the numerical procedures. I used both to evaluate model performance predicting ozone. I relied primarily on graphical data assess model performance predicting observed precursor concentrations. The several procedures utilized appear to lead to consistent conclusions.

7. Does the model/mechanism meet U.S. EPA and ARB model performance acceptance criteria?

With respect to ozone, all of the models meet these criteria (gross error, normalized bias, peak ratios) in most of the zones most of the time. I found however, that these criteria were among the easiest to meet of those that I used to assess performance. Based on a more complete analysis, I believe the CALGRID/SAPRC99f model performs the best in predicting ozone. There are no criteria for assessing performance predicting NO or NO2. Based primarily on a review of graphical material, CAMx/SAPRC99f appears to perform a little better than CALGRID/SAPRC99f predicting these precursors.

8. Considering all of the factors above, is this model/mechanism significantly better than the others?

As a result of my performance review, past scientific critiques and practical considerations, I believe the CALGRID/SAPRC99f model should be the preferred approach to support the 2003 SIP revision.

Three additional general questions have been asked by the AQMD.

9. Using the existing tools and information, is there anything that could be done to provide greater confidence in the modeling used in the AQMP?

Yes. I believe a data base exists or can be created relatively easily which would enable the AQMD to compare the *response* of predicted and observed ozone to changes in emissions and/or changes in emissions + meteorology. Such an analysis should be done. Ideally, performance of model response could be evaluated by reconstructing 1987 emissions using updated estimation procedures. Observations and predictions could then be noted for the 1987 episodes. Next, 1997 emissions and episodes could be run. Changes in predictions vs. changes in observations could then be compared. These changes would be attributable to differences both in emissions and meteorology. Another, more subjective, approach might be used to assess the accuracy of a model's response to primarily to changes in emissions. If 1997 episodes "similar" to the 1987 episodes could be selected, it would be possible to assess accuracy of a model's response that is due, primarily, to changes in emissions. Evaluations of this sort would help circumvent the concern expressed by some about "getting the right answers for the wrong reasons". The reason some feel this concern is valid, is that when model performance is evaluated solely on the ability to replicate one or more past episodes, one is not asking the question which is of paramount concern: how accurately does a model predict *changes* in air quality? If one asks this key question directly in the performance evaluation, getting the right answer provides greater confidence than asking a related, but less relevant question for SIP purposes. I would like to see this test applied for both the CALGRID/SAPRC99f and CAMx/SAPRC99f models.

10. Many of the model runs exhibit underestimation of observed peaks to some degree. Is there sufficient understanding of the reason(s) for the underestimation to warrant making an adjustment?

I do not believe there is, at least for this SIP revision. The process analysis technique, which I believe can be utilized with the CAMx model, might provide some useful subsequent insights about appropriate changes. I also think it would be useful to apply the CALGRID/SAPRC99f, CAMx/SAPRC99f and UAM/CB4 approaches and then use the relative reduction approach with the appropriate ozone design values for several of the zones. This would likely provide some insight into the potential significance of using absolute predictions of ozone in an attainment demonstration with a model that systematically underpredicts observed peak hourly concentrations of ozone.

11. Is model performance acceptance biased by the underestimation of the peak concentrations?

Ability to predict peak observed ozone concentrations is a very important consideration. Ratio of peak predicted to peak observed concentrations is one of three U.S. EPA ozone performance measures for which a criterion exists. Predicted future peak concentrations are generally compared to the level specified in the ozone NAAQS to judge whether a SIP is adequate. As I noted in Section 2.0, I try to consider how model results are typically used in a demonstration in order to judge how to evaluate performance of the 5 models under consideration. Thus, several measures related to the ability to predict the magnitude of peaks are included in my assessment. There are other measures, such as timing, ability to predict exceedances, normalized bias and gross error that are also considered.

One should remember that no matter how scientifically current certain components of a model are, what is also critical is how these components interact to produce an air quality prediction. Regardless of how sophisticated a model is, assumptions have to be made about how to consider emissions, parameterize meteorological and chemical phenomena, etc., etc. The best combination of assumptions is not intuitively obvious. This is why the ability to produce an answer that appears to coincide with observations is a very important consideration in choosing a model for use in a SIP.

In short, I believe it is appropriate to give heavy consideration to a model's ability to predict a measure that is a key one in accepting an attainment demonstration. I would not characterize this as "bias".

4.0 Conclusions About PM10 and PM2.5

4.1 UAM-LT Predictions and Observations of PM10

Annual and quarterly mean predictions and observations are provided for the following components of PM: ammonium (NH4), nitrate (NO3), sulfate (SO4), organic carbon (OC), elemental carbon (EC) and "Other". Comparisons are an aggregate of days when 24-hour mean observations are taken, typically once in 6 or once in 3 days. AQMD has defined acceptable model performance for each component as agreement between prediction/observation spatially averaged pairs to within $\pm$ 30%. The basis for this criterion is unclear, but I have nothing better to offer.

Looking at comparisons between spatially averaged observations and predictions for annual means of each component, the performance criterion is satisfied for NH4, NO3 and OC components, or about 86% of the reported mass (i.e., the sum of the 6 components). Measurements are made at 5 sites. Acceptable performance for spatially averaged annual concentrations, as well as agreement within $\pm$ 30% at individual sites, is as follows:

Anaheim---NH4, NO3, OC, EC
Diamond Bar---NH4, NO3, OC
Fontana---NH4, NO3, SO4, Other
LA---NH4, NO3, OC
Rubidoux---NH4, NO3, SO4, OC, EC, Other
Spatial Average---NH4, NO3, OC

However, due to seasonal variations in emissions and meteorology, it is good practice to evaluate performance (as well as perform strategy simulations) separately for each quarter of the year. Upon doing this, one finds generally poor performance for UAM-LT during the first quarter. The model typically overpredicts observations during this quarter. Indeed, agreement within $\pm$ 30% is limited to the following during quarter 1:

Anaheim---EC
Diamond Bar---SO4, OC, EC
Fontana---NH4, NO3, OC, Other
LA---SO4, EC
Rubidoux---none.
Spatial Average---EC (~7% of the observed spatially averaged mass of PM10 is predicted "correctly")

Of the 4 quarters, observed mean PM10 concentrations are generally lowest during the first quarter.

Performance improves for the second quarter. Agreement within $\pm$ 30% is noted below.

Anaheim---NH4, NO3, OC, EC, Other
Diamond Bar---OC, EC, Other
Fontana---NH4, NO3, OC, Other
LA---NH4, NO3, OC, EC, Other
Rubidoux---OC, EC, Other
Spatial Average---NH4, NO3, OC, EC, Other (~90% of the observed spatially averaged mass of PM10 is predicted "correctly")

Observed quarterly means for the second quarter are generally not as high as those observed in the third and fourth quarters.

Agreement within $\pm$ 30% during the third quarter is as follows:

Anaheim: NH4, NO3, OC, EC
Diamond Bar---NH4, NO3
Fontana---SO4, Other
LA---NH4, NO3, OC
Rubidoux---NH4, NO3, Other
Spatial Average---NH4, NO3, Other (~64% of the observed spatially averaged mass of PM10 is predicted "correctly")

Observations during the fourth quarter are generally 50-60% higher than those seen during the next highest (i.e., the third) quarter. Agreement within $\pm$ 30% during the 4th quarter occurs as follows:

Anaheim---NH4, NO3, OC
Diamond Bar---NO3, OC
Fontana---SO4, Other
LA---NO3, OC
Rubidoux---SO4, OC, Other
Spatial Average---NO3, OC (~49% of the observed spatially averaged mass of PM10 is predicted "correctly")

Looking at projections to 2006 and 2010, it appears the annual NAAQS will not be met at only 3 of 57 sites. Further, the projected violations are within about 10% of the concentration level specified in the NAAQS. As noted below, it is projected to be much more difficult to meet the annual NAAQS for PM2.5. Thus, once a SIP is developed for PM2.5, attainment of the PM10 NAAQS should "follow along" from efforts to meet the PM2.5 NAAQS. Further, if efforts to meet the PM10 NAAQS focus on reducing the "Other" (primary?) component, they will have little impact on reducing PM2.5.

4.2 UAM-LT Predictions and Observations of Annual/Quarterly PM2.5

Although AQMD is not yet required to submit a SIP revision addressing PM2.5, a modeling analysis has been performed in anticipation of a future requirement to do so. The same performance criteria are applied to the same components noted in the discussion of PM10. The extent to which mean quarterly or annual predictions agree with corresponding observations within ± 30 % is noted below.

Annual Concentrations

Anaheim---NH4, NO3, OC
Diamond Bar---NH4, NO3
Fontana---NH4, NO3
LA---NH4, NO3, OC
Rubidoux---SO4, OC, EC
Spatial Average---NH4, NO3, OC (~73% of the observed PM2.5 is predicted "correctly")

Note that for the spatial average, it is 73% of the *measured* mass has associated acceptable model performance.

Quarter 1

Anaheim---SO4, EC
Diamond Bar---NH4, NO3, SO4, EC
Fontana---NH4, NO3, EC
LA---NH4, SO4, OC
Rubidoux---None
Spatial Average---EC (~11% of the observed PM2.5 is predicted "correctly")

Quarter 2

Anaheim---NH4, OC, EC
Diamond Bar---NH4, OC
Fontana---NH4, NO3
LA---NH4, OC
Rubidoux---OC, EC
Spatial Average---NH4, NO3, OC, EC (~72% of the observed PM2.5 is predicted "correctly")

Quarter 3

Anaheim---NH4
Diamond Bar---NH4
Fontana---NH4
LA---NH4
Rubidoux---NH4, NO3
Spatial Average---NH4 (~14% of the observed PM2.5 is predicted "correctly")

Quarter 4

Anaheim---NO3, OC
Diamond Bar---NO3, OC
Fontana---SO4, OC
LA---OC
Rubidoux---SO4, OC
Spatial Average---OC (~18% of the observed PM2.5 is predicted "correctly")

The seasonal pattern for observed PM2.5 is similar to that for PM10, only more pronounced. Quarterly average values during the 4th quarter are about double that of the next highest quarter. This is mostly attributable to large increases in nitrate and ammonium.

One major difference between model performance for PM2.5 and PM10 is the very poor performance predicting the "Other" component for PM2.5. The model drastically overpredicts this component (by factors ranging from "3" to "61"). Although the "Other" component is the least important component according to the observations, it is the second (to nitrates) most important component according to the model's predictions. Thus, the potential exists for misleading results when simulating strategies. Generally poor performance in predicting PM2.5 as well as the relative importance of components of PM2.5 argues for using a "relative reduction factor" (RRF) type of approach for PM2.5 strategy simulations. Note also that the comparisons emphasize the need to consider each quarter separately. Looking simply at the annual averages implies better performance than is actually the case.

As with PM10, it appears that sulfates and all, or nearly all, of the nitrates are likely to be ammonium salts.

Simulations projected to 2006 and 2010 suggest that meeting the annual NAAQS for PM2.5 will be a challenge in the AQMD. 36 of 57 sites have projections exceeding the annual NAAQS. Assuming PM2.5 is the sum of the 6 components which are reported, levels exceed that of the annual NAAQS by anywhere from 0.1 to 12.1 micrograms/cubic meter.

4.3 Some Concerns with the Analysis

Details of the modeling analysis for particulate matter (and for that matter about the model which was used) are sketchy. Thus, it is not possible to comment at length about what was done. However, based on the material reviewed, I have at least two concerns.

First, it appears that an assumption has been made that the spatially averaged composition observed at 5 monitoring sites is representative of the entire modeling domain. Predicted differences in each of the components were then most likely weighted according to their importance to derive future concentrations of PM10 and PM2.5 at 50-60 monitoring sites. If this was what was done, a great deal more effort is needed to explain the procedure and to justify the use of just 5 locations to draw inferences about the entire domain.

The second concern is related to the first. It has to do with using spatially averaged predicted and observed concentrations to evaluate model performance. Due to the incommensurability problem, I think use of spatial averages may be a legitimate way to assess a grid model's performance predicting concentrations of primary pollutants. However, it seems to me that the evaluation should be preceded by some effort to document why it is appropriate to group information from certain sites, be it proximity, presence of nearby, similar sources, or whatever.

Since the PM10 problem in the AQMD appears to be a relatively minor one, a great deal more in the way of a modeling effort may not be warranted. However, PM2.5 looks to be a major future problem. Draft U.S. EPA guidance for modeling PM2.5 (U.S. EPA, 2001) closely parallels the Agency's draft guidance for applying models to address the 8-hour NAAQS for ozone. Both sets of guidance require developing a modeling protocol and addressing many of the same problems which the AQMD faced in applying models to address the 1-hour NAAQS for ozone. The guidance recommends using models in a relative sense to develop relative reduction factors that are applied to observed air quality.

5.0 Recommendations for Additional Work

1. AQMD should document what modeling assumptions, modeling preprocessors and/or components were *actually used* to produce the estimates that underlie the strategy reflected in the 2003 SIP revision. It is clear that the approach originally outlined in the modeling protocol has been changed in response to subsequent findings and discussions.

2. As another peer reviewer has suggested, AQMD needs to put a great deal of effort into producing a relatively brief (e.g., 15-20 page) narrative summarizing what analyses were done and how they lead to the conclusion that the SIP revision will meet its intended goals. This narrative should be aimed at managers and the lay public. More detailed descriptions (e.g., for CARB and U.S. EPA technical staffs) should be in appendices.

3. I strongly recommend that model performance in predicting secondary pollutants like ozone, NO2 and particulate mass associated with nitrates, sulfates and ammonium ion focus on how accurately the model is able to replicate the observed *response* to changes in emissions, meteorology and both. In the near future, this can likely only be done for ozone and its precursors. I recommend that AQMD revise 1987 emissions (using the current, state of the art methodologies), rerun the 1987 episodes with the revised emissions and compare changes in predicted ozone and NO2 between the 1987 and 1997 episodes with the observed changes between the 1987 and 1997 episodes. Adding this approach to the others for model performance evaluation, addresses a key concern about model performance: how well does the model replicate the *response* of secondary pollutants to changes in precursor emissions and/or changes in meteorology. Contrasting changes in predictions vs. observations for weekends and weekdays is another potential way to evaluate a model's response. However, I am less sure whether the differences in emissions on weekends vs. weekdays are sufficient to send a signal strong enough to be discerned over "noise" attributable to uncertainties in weekend vs. weekday inventories.

4. AQMD should retain and archive all files that are needed to perform simulations in the future that contrast predicted vs. observed changes in ozone over time. These files should also be made available to the research community to test whether incremental changes (reflecting more current scientific findings) improve the performance of grid models.

5. Some additional suggestions about model performance evaluation for predicting ozone and its precursors follow.

(a) Site by site comparisons in which predictions within a surface grid cell are matched with observations occurring within the cell at the same time may present a test that is needlessly restrictive. For example, the EPA Guidance for the 1-hr ozone NAAQS (pp.10-11) (U.S. EPA, 1996) suggests looking at a 3x3 array of 5-km cells centered on the monitor and using the highest prediction within this array in what that guidance defines as "the statistical attainment test".

U.S. EPA monitoring regulations suggest that ozone monitors should represent an "urban scale". This is a rather wide range, but the modeling guidance has selected a value (15 km) near the lower end of the range. The recommendation to consider an array of cells was also made in recognition that, in many cases, it may be nearly impossible to get the exact time and location correctly and that, in any event, small discrepancies may not be important. Likewise, it may often be legitimate to compare the highest nearby prediction within $\pm$ 1-2 hours to each site's peak observation when evaluating model performance, to see whether a seemingly poor result is simply a reflection of a bunch of "near misses".

(b) The time series correlations for predicted vs. observed ozone presented by AQMD are almost uniformly bad. However, this may reflect the relatively narrow range of observations available at most sites on the days selected for modeling. Excluding all hours with ozone observations <.06 ppm from the analysis exacerbates this problem. Often, the range of observations is only .06 – about .13 ppm. If these low correlations are a concern, AQMD should repeat the analysis including all observations and predictions to increase the range.

(c) Comparisons may yield worse than expected results due to the apparently different rounding conventions used for the monitoring data (to the nearest whole pphm for everything but CO) and modeled data, reported to the nearest 0.1 pphm (for everything but CO). Differing conventions for reported monitoring and modeled data may be even more of a problem for CO, as it appears measured values of this pollutant are rounded to the nearest 0.5 ppm. Much of the diurnal data presented for CO show little hour-to-hour variability, in part perhaps, due to this convention. If it is possible to access more precise observed data from the archives, AQMD should see whether using identical levels of precision in the predicted and observed data improves performance.

(d) I urge AQMD to make greater use of special air quality measurements that may have been taken during the SCOS97 field program to help evaluate model performance. More specifically, observed and predicted "indicator ratios" of pollutants should be looked at when the measurements permit. Accurate prediction of observed indicator ratios may suggest that a model's ozone predictions will respond correctly to changes in precursor emissions. Identity and use of indicator ratios are described in Sillman (1995), Sillman (1998) and Blanchard, *et al.* (2000), as well as elsewhere.

6. Not surprisingly, agreement between predicted and observed primary pollutants (NO and CO) is poor. In my opinion, this likely reflects an incompatibility between assuming a uniform concentration of these pollutants within a 5-km grid cell and effects of smaller scale phenomena on the monitored observations. It may be preferable to average modeled and monitored data of these pollutants over several sites when making comparisons. Perhaps this might help random, small-scale fluctuations in the observations to balance. Another potentially useful approach may be to compare observed and predicted *ratios* of primary pollutants.

If spatial averaging is used, there should be an accompanying rationale for combining the data from several locations. This was lacking in the UAM-LT performance evaluation comparing observed and predicted components of PM10 and PM2.5.

7. I believe its underlying science is sufficient and it performed sufficiently well for the recommended primary model for ozone (CALGRID/SAPRC99f) to be used to support the 2003 SIP revision for ozone.

Nevertheless, performance of CALGRID/SAPRC99f is not outstanding. Further, my assessment is based on only 3 modeled days. It is also troubling that CAMx/SAPRC99f and UAM/CB4 appear to better replicate observed concentrations of NO2. Therefore, I believe AQMD should make plans to perform a "mid-course review" in accordance with U.S. EPA guidance (U.S. EPA, 2002), as well as perform a weight of evidence (WOE) analysis. More specifically, AQMD should archive modeling files used to create predictions for the 1997 episodes. Project emssions to a future "mid-course" date (e.g., 2006) and to the attainment date (e.g., 2010) and apply the model to obtain corresponding estimates. When 2006 comes, derive ambient trend information and normalize this for meteorological differences. Compare the apparent relative progress toward the NAAQS from the normalized trend data against the relative progress projected for 2006 by the model. By "relative progress", I mean the percentage of the way between 1997 predicted (or observed) air quality and the air quality goal (e.g., 124 ppb) that is predicted (or observed) in 2006. If the relative change in the air quality observations is less than that predicted, a SIP adjustment or reexamination of the model may be needed.

The weight of evidence analysis could include looking at changes predicted between 1997, 2006 and 2010 with the most current model in 2006 that exhibits adequate performance. Whether a new SIP revision is deemed necessary may depend on whether the normalized trend in observed ozone leads to a relative changes in high ozone which is less than that predicted with the updated model.

8. Because it appears likely that the PM10 NAAQS will be close to being attained and subsequent efforts to meet the NAAQS for PM2.5 will likely lead to improvement in PM10 as well, I do not believe that a major additional effort to model PM10 is necessary. AQMD should instead focus its efforts on performing modeling to support a SIP to meet the NAAQS for PM2.5. As outlined in draft U.S. EPA modeling guidance related to the PM2.5 NAAQS (U.S. EPA, 2001), this will require a major effort, comparable to the ones needed to support the 1-hour and 8-hour NAAQS for ozone.

6.0 Summary

In this review, I first examine the ability of 5 modeling approaches to replicate observed ozone concentrations during a 3-day episode in 1997. In Section 2.1, I use six measures to evaluate the models' performance in each of 9 geographical zones identified by the AQMD. Choice of measures is influenced by my familiarity with how modeling results are used by regulatory agencies to determine whether a SIP is likely to lead to attainment of the NAAQS for ozone. Using these measures, CALGRID/SAPRC99f performs substantially better than the other models. CAMx/SAPRC99f performs second best. UAM/CB4, CALGRID/CB4 and CAMx/CB4 perform comparably.

In Section 2.2, I review performance of CALGRID/SAPRC99f, CAMx/SAPRC99f and UAM/CB4 in predicting observed concentrations of NO, CO and NO2. Evaluation of this performance is based on graphical presentations and is more subjective than the evaluation for ozone. For understandable reasons, no model does very well predicting NO or CO. CAMx/SAPRC99f and UAM/CB4 do better than CALGRID/SAPRC99f in replicating observed concentrations of NO2. The latter model's predictions of NO2 are lower than those of the former two on some days at some sites.

Despite the outcome of the NO2 comparisons, in Section 2.3 I recommend that AQMD use CALGRID/SAPRC99f as the primary modeling approach to support the 2003 SIP revisions for ozone. This recommendation reflects greater confidence in the observed ozone data, larger differences among the models than was true for NO2, as well as some practical considerations.

In Section 3.0, I address a series of questions posed by the AQMD. In general, these questions concern (a) whether the five ozone models considered for use by the AQMD are sufficiently well grounded scientifically; (b) whether the methodologies used by the AQMD in applying the models are adequate, and (c) whether the procedures followed are consistent with regulatory guidance issued by the CARB and the U.S. EPA. My sense is that all of the models are adequate provided those applying them prepare a detailed protocol which is discussed in the local scientific community, model performance is assessed and found adequate and diagnostic tests are applied to see how sensitive conclusions might be to assumptions which need to be made in the modeling.
In general, the assumptions that were made by the AQMD and CARB in applying the models seem plausible. Finally, I did not detect major inconsistencies with U.S. EPA guidance nor problems relative to the guidance that cannot be easily addressed.

In Section 4.0, I summarize some of the findings regarding application of the UAM-LT model to estimate annual and quarterly mean concentrations of PM10, PM2.5 and 6 components of these mixtures. In general, model performance is pretty good for PM10. Typically, components which comprise well over 50% of the measured mass of PM10 are predicted accurately according to the criterion of $\pm$ 30% agreement established by the AQMD. The exception occurs in the winter, when performance is poor, but observed concentrations are low. Usually, the relative amounts that each component comprises of predicted spatially averaged PM10 agree reasonably well with spatially averaged observations.

Performance predicting PM2.5 is not good. Usually, less than 50% of the measured mass is accurately predicted. As with PM10, performance is worst during the winter months. More significantly, performance is not good during the 3rd and 4th quarters, when observations are the highest. More significantly still, the model substantially overestimates the relative importance of primary particulate matter. This latter shortcoming could lead to selection of ineffective control strategies, unless models are applied in a relative sense using relative reduction factors.

The documentation for the PM10 modeling is poor and should be improved. However, I do not believe substantial additional resources need to be spent on PM10 modeling. I recommend modeling of particulate matter focus on PM2.5 and that results should be used in a relative sense using relative reduction factors.

Section 5.0 contains a series of recommendations concerning documentation, performance evaluation, data archiving, mid-course reviews and use of weight of evidence in subsequent analyses. Key recommendations follow.

-A major effort should be made to summarize for managers and the lay public the modeling used to justify the strategy reflected in the SIP revision.

-More emphasis should be placed on efforts to evaluate how well the model responds to *changes* in emissions and meteorological conditions.

-Although the available models provide an adequate basis for selecting SIP control strategies, there are uncertainties associated with the predictions. There should be ongoing efforts to perform a mid-course review which includes comparing the response of the chosen modeling approach to changes in emissions and meteorology to that obtained with the most current, viable model (e.g., in 2006) and with observed air quality trends normalized for meteorological fluctuations.

-Care should be taken to archive modeling files to facilitate future mid-course reviews, performance evaluations and model improvements.

7.0 Cited References

Blanchard, C.L., P.M. Roth, S.J. Tanenbaum, S.D. Ziman and J.H. Seinfeld, (2000), "The Use of Ambient Measurements to Identify Which Precursor Species Limit Aerosol Nitrate Formation", Accepted for publication in *J.Air & Waste Management Association*.

Sillman, S., (1995), "The Use of N0y, H2O2 and HNO3 as Indicators for O3-N0x-ROG Sensitivity in Urban Locations", *J.Geophys. Res. 100,* pp.14175-14188.

Sillman, S., (1998), *Evaluating the Relation Between Ozone, N0x and Hydrocarbons: The Method of Photochemical Indicators*, EPA/600R-98/022.

U.S. EPA, (1991), *Guideline for Regulatory Application of the Urban Airshed Model*, EPA-450/4-91-013.

U.S. EPA, (1996), *Guidance on Use of Modeled Results to Demonstrate Attainment of the Ozone NAAQS*, EPA-454/B-95-007.

U.S. EPA, (1999), *Draft Guidance on the Use of Models and Other Analyses in Attainment Demonstrations for the 8-Hour Ozone NAAQS*, EPA-454/R-99-004.

U.S. EPA, (2001), *Guidance for Demonstrating Attainment of Air Quality Goals for PM2.5 and Regional Haze (draft 2.1)*, January 2, 2001, available on www.epa.gov/ttn/scram.

U.S. EPA, (2002), *Recommended Approach for Performing Mid-course Review of SIP's to Meet the 1-hour NAAQS for Ozone*, January 2001, available on www.epa.gov/ttn/scram.

This page is left blank intentionally

TO: Barry Wallerstein
Joe Cassmassi

FROM: John Seinfeld

DATE: January 21, 2003

SUBJECT: Initial Critique of AQMP Models and Results

Five models have been used to simulate the August 4-6, 1997 episode, for which the peak 1-hour ozone level is 187 ppb, which occurred on August 5 at Riverside. The ΔVOC is that needed to reduce the 1-hour peak ozone from 187 to 120 ppb, a ΔO_3 of 67 ppb. If a model exactly predicts the 187 ppb peak concentration for August 5, 1997, then one needs only to vary VOC emissions until the peak is reduced to 120 ppb. The level of VOC emissions August 4-6, 1997 for the meteorological conditions of required to lower the peak to 120 ppb then becomes the so-called "VOC carrying capacity" of the Basin. The actual ΔVOC that would be required in 2010 is the different between the estimated base VOC emissions in 2010 (about 600 tons/day) and the calculated VOC carrying capacity.

If a model does not reproduce the peak 187 ppb on August 5, the ΔO_3 required to reach 120 ppb is different from 67 ppb. Those models that do not predict 187 ppb all happen to underpredict the peak value. A typical peak prediction might be 160 ppb. Then the ΔO_3 required from that predicted peak is only 40 ppb, rather than the actual 67 ppb. Consequently, the VOC carrying capacity that results from that model will reflect an ozone reduction of only 40 ppb and cannot be expected to accurately reflect the ΔVOC needed if starting from a 187 ppb peak O_3. So, if a model is used that underpredicts the peak O_3 concentration on the design day, one issue that must be dealt with is how to correct the VOC carrying capacity to reflect the effect of that underprediction.

The elements of a model simulation include:

- emissions
- meteorology (velocity and temperature field, mixing depth)
- chemistry
- boundary conditions
- dry deposition
- vertical and horizontal diffusion
- numerical analysis of advection, diffusion, and chemical reaction
- grid system

Each of these elements is complicated in its own right. Ideally, when comparing performance of several models, one would like to have identical emissions meteorological fields, and boundary conditions, so that differences in prediction reflect differences in chemical mechanisms, dry deposition treatments, diffusion treatments, and numerical methods used to solve the governing equations. Even so, interactions among these submodels are nonlinear so that behavior of one submodel can compensate for inaccuracies in another submodel to produce a simulation that matches actual data. The difficulty with this situation is that when emissions are reduced these factors that compensate each other in the base case may not do so in the reduced emission case.

In principal, one should employ the model that most accurately reflects the underlying physics and chemistry. Of the five models, UAM, and especially CB4, is known to be the most out of date. UAM also has a notoriously inaccurate advection solver (Smolarkewicz) that strongly smears out concentration distributions. Yet, the UAM CB4 simulation of the August 4-6, 1997 episode produces the closest match to the observed peak O_3 of the five models. Enhanced photochemical production of O_3 is, most likely, compensating for the artificial spreading induced by the Smolarkewicz advection algorithm. SAPRC-99 is the preferred chemical mechanism, and indeed would be considered as state-of-the-science. Both CALGRID and CAMx represent improvements over UAM in scientific treatments; however, both of these models underpredict peak O_3 on August 5, 1997 with both SAPRC 99 and CB4. If CB4 is "hot" chemically, as suggested above, one might expect O_3 predictions to be higher with both CALGRID and CAMx using CB4; this is not the case.

I have not been associated with this AQMP modeling process long enough to be able to diagnose why UAM is matching the design day peak O_3 and why the other models are underpredicting. The AQMD staff should continue to attempt to determine the underlying explanations for the behavior of the models through diagnostic simulations, and I will assist with designing these as much as possible. In summary, however, it is not apparent why each of the models is performing as they are. As discussed at the meeting at the AQMD, based on overprediction of NO_x levels in the western portion of the Basin, it appears that there may be some problem with CAMx.

While the AQMD staff should continue to identify the reasons for underprediction of the peak O_3 by each of the four models, the AQMD must decide on a strategy for determining the VOC carrying capacity of the Basin in 2010 based on the meteorology of the August 4-6, 1997 episode in the event that it is not possible to improve the simulations before the AQMP must be submitted. The following potential strategies exist:

1. Because it is close to the actual observed peak O_3, use the UAM CB4 simulation as the basis for determining the VOC carrying capacity.
2. Use either the CALGRID or CAMx simulations as the basis for determining the VOC carrying capacity with no adjustment for underpredicted peak O_3.
3. Artificially increase the CALGRID or CAMx predicted peak O_3, perhaps by enhancing photolysis rates, to match the measured O_3 and use that simulation as the base case.
4. Determine the percentage underprediction of peak O_3 in the base case simulation and reduce the target concentration of 120 ppb by the same percentage.
5. Starting from the simulated peak O_3, determine the VOC reduction required for a reduction of ΔO_3 = 67 ppb.

None of these is ideal. In the absence of more information, I would recommend option #4. By reducing the target concentration (120 ppb) by the percentage by which the peak O_3 is underpredicted, one obtains a more realistic ΔO_3 from which to determine the VOC carrying capacity. Hopefully, it will be possible to continue to evaluate the model simulations over the next couple of months.

MEMORANDUM

DATE: January 13, 2003

TO: Joe Cassmassi

FROM: Mel Zeldin

SUBJECT: Ozone Modeling Evaluation and Suggestions

This memo is in response to your request for written input regarding the ozone modeling evaluation as an outcome of the meeting held last Friday. As I stated in the meeting, it is my belief that the UAM model has been the historical workhorse for ozone attainment modeling purposes for the last decade, and in order for another model to be used, either of two conditions would need to occur: (1) the UAM model would have to be shown to be deficient in the validation process; or (2) either the CALGRID or CAMx models would have to be shown to be definitively superior to the UAM.

In my opinion, neither of these conditions is evident in the validation statistics. While there are advantages and disadvantages shown for each model, there is clearly nothing to show that any one model is superior to the others. Therefore, I would suggest that the UAM continue to be the model used for the ozone attainment process. I believe that other members of the Review Panel shared that opinion. I further believe that changing to another model, without clear superiority of that model, could be considered "gaming" or "model shopping" to a desired result.

In my advisory capacity to the Staff, however, I do have some suggestions:

1) The AQMP (Plan) should be based on the UAM results at this time.

2) It should be noted in the Plan, however, that the science experts believe that, in theory, at least, the CALGRID and CAMx models utilize newer techniques that should be better than the older UAM. It should be further noted that the ARB believes the CALGRID model to be the preferred model.

3) It should further be noted that due to the number of model runs involved, and the short time frame in which to produce the attainment demonstration to avoid conformity lapse, one episode is deemed to be insufficient to determine if either CALGRID or CAMx are superior to the UAM.

4) The District should commit to perform additional model evaluations of other episodes to determine if indeed CALGRID or CAMx show superior performance to the UAM, and these evaluations would be completed within a two-year time frame.

5) Since the UAM has the lowest VOC carrying capacity, and it is presumed that the additional tonnages to be reduced would be placed in the "black box," the District would commit not only to the larger "black box," but also to proceed with studies, in parallel during the same two-year period, to explore and evaluate advanced technologies for additional VOC reductions. Any rulemaking activity to gain the extra VOC tonnages would be deferred until 2005 – when the model evaluation would be completed.

6) As part of the current Plan and SIP submittal, the District should state that, at the end of the two-year period, the results of the model evaluations will be presented. If the results do not show any superiority from either CALGRID or CAMx, then the District would proceed to move forward with identifying measures in the "black box" to get to the UAM VOC carrying capacity. If, on the other hand, the additional model evaluations show that either the CALGRID or CAMx models are superior to the UAM, then the extra tonnages assigned to the "black box" would be removed. This would likely involve a SIP amendment.

7) In taking this course of action, the District should state that there is too much uncertainty in the model evaluations at this time, and is therefore erring on the side of public health interests until a more complete evaluation is completed. In stating this, the District should make clear at this time, that any changes to the carrying capacity, at the end of the two-year period, are reflecting better science in evaluating future year attainment, and removing "black box" tonnages are not to be considered "backsliding," since attainment would be demonstrated with a (presumably) higher VOC carrying capacity. If either CALGRID or CAMx were shown to be the superior model, the only thing "lost" by the District would be the parallel effort for identifying new VOC control technologies, and I don't think such an effort could ever be deemed "wasted" in the sense that such methods may ultimately be needed for 8-hr attainment in the future.

8) In making additional model evaluations, I would suggest using the other episodes mentioned in the model protocol: September 26-29, 1997 and July 13-18, 1998. While only the August 1997 episode is used for attainment demonstration purposes, I see no reason why the other episodes can't be used for model evaluation purposes only. Since some effort has been made to identify these episodes already, presumably these events will be easiest to develop the model inputs. Also, I would recommend that the August 1997 episode be "backcasted" to 1987 for an additional evaluation when emissions were considerably different than the 1997 period. An evaluation of an episode in 1987, which is meteorologically similar to the 1997 episode, could be used for approximated validation data, or, as you suggested, a mean value for that meteorological classification could be used. This will at least demonstrate if any of the new models "blow up" with a 1987 inventory, or are able to reasonably simulate those conditions.

9) I like the idea suggested by another panel member for some kind of scoring system for evaluating model performance, other than just the performance statistics. This could be something that evaluates: the pattern representativeness (i.e, the modeled ozone spatial pattern replicates the observed pattern); the peak ozone representativeness; the temporal representativeness (i.e., the timing of the peaks at sites across the Basin match the observed time of the peaks); and similar evaluations for other pollutants, such as NO and NOx. It would be advisable to establish a scoring system prior to undertaking the additional model evaluations.

I hope these comments are helpful. Please call me if you have any questions or would like to discuss further.

ATTACHMENT 3

Mid-Course Modeling Reviews

ATTACHMENT 4

CEPA Source Level Emissions Reduction Summary for 2006: Annual Average Inventory

Year 2006 Emission Reductions Excluding Natural Sources by Control Measure in the South Coast Air Basin
(Annual Average Inventory - Tons/Day)

(A) Reductions Without Overlapping/Double-Counting With Other Control Measures (1)

		(Reductions - Tons/Day)					
Measure	Name	VOC	NOx	CO	SOx	PM10	PM2.5
BA-2202	Baseline adjustment for R2202	2.48	2.80	29.68	0.02	0.06	0.06
BA-1122	Baseline adjustment for R1122	3.01	0.00	0.00	0.00	0.00	0.00
BA-1113	Baseline adjustment for R1113B	-0.01	0.00	0.00	0.00	0.00	0.00
BA-1133	Baseline adjustment for Composting Opr.	-3.68	0.00	0.00	0.00	0.00	0.00
BA-FVR	ARB-Petroleum Marketing - Bas. Adj.	-1.20	0.00	0.00	0.00	0.00	0.00
BA-ARPT	ARB-Airport Ground Support Equip - Bas. Adj.	-0.22	-1.17	0.00	0.00	0.00	0.00
BA-SMOG	ARB-Smog Check II - Bas. Adj.	2.54	4.53	0.00	0.00	0.00	0.00
CTS-07	FER Architectural Ctg (R1113 Phase 3) (VOC)	3.27	0.00	0.00	0.00	0.00	0.00
CTS-10	Misc. Industrial Coating&Solvent Opr. (VOC)	0.86	0.00	0.00	0.00	0.00	0.00
FUG-05	ER from fugitive Emission Sources (VOC)	1.67	0.00	0.00	0.00	0.00	0.00
CMB-09	Petroleum FCCU (PM10,PM2.5,NH3)	0.00	0.00	0.00	0.00	0.09	0.08
BCM-7	FER from Fugitive Dust Sources (PM10)	0.00	0.00	0.00	0.00	0.00	0.00
BCM-08	Aggregate Operation (PM10)	0.00	0.00	0.00	0.00	0.62	0.33
PRC-03	COE fr Restaurant Operations (PM10)	0.00	0.00	0.00	0.00	0.33	0.33
PRC-07	Industrial Process Operations (VOC)	0.81	0.00	0.00	0.00	0.00	0.00
WST-01	Ems Reduction fr Livestock Waste (VOC,NH3)	4.81	0.00	0.00	0.00	0.00	0.00
WST-02	COE fr Composting (PR1133) (VOC,NH3,PM10)	0.40	0.00	0.00	0.00	0.00	0.00
CMB-07	Refinery Flares (VOC,NOX,SOX,CO,PM)	0.17	0.17	0.99	2.16	0.03	0.03
Grand Total (Net)		14.92	6.32	30.68	2.18	1.13	0.83

EMISSION SUMMARY FOR
(POINT, AREA, MOBILE SOURCE, AND OFF-ROAD MV)

BASELINE EMISSIONS	VOC	NOx	CO	SOx	PM10	PM2.5
Point source	66.68	13.96	41.56	5.95	12.39	11.06
Area source	229.20	49.09	172.45	0.96	242.95	70.37
RECLAIM	0.00	34.20	0.00	12.03	0.00	0.00
Total Stationary	295.88	97.25	214.00	18.94	255.34	81.43
On-road	251.82	548.18	2610.49	4.70	18.03	12.02
Off-road	105.85	280.49	812.80	34.04	19.25	17.07
TOTAL	653.55	925.92	3637.29	57.67	292.63	110.51
EMISSION REDUCTIONS						
Point source	2.16	0.17	0.99	2.16	0.40	0.28
Area source	7.96	0.00	0.00	0.00	0.67	0.49
Total Stationary	10.12	0.17	0.99	2.16	1.07	0.77
On-road	5.02	7.33	29.68	0.02	0.06	0.06
Off-road	-0.22	-1.17	0.00	0.00	0.00	0.00
TOTAL	14.92	6.32	30.68	2.18	1.13	0.83
REMAINING EMISSIONS						
Point source	64.52	13.80	40.57	3.79	11.99	10.79
Area source	221.24	49.09	172.45	0.96	242.29	69.88
RECLAIM	0.00	34.20	0.00	12.03	0.00	0.00
Total Stationary	285.76	97.09	213.01	16.77	254.27	80.66
On-road	246.80	540.85	2580.81	4.68	17.97	11.96
Off-road	106.07	281.66	812.80	34.04	19.25	17.07
TOTAL	638.63	919.60	3606.62	55.49	291.50	109.68
NSR/Set-Aside	0.00	0.00	0.00	0.00	0.00	0.00
GRAND TOTAL (T/D) (2)	638.63	919.60	3606.62	55.49	291.50	109.68
Mobility Adjustments (3)	2.25	1.29	37.59	0.02	0.24	0.17

(1) Emission reductions for individual measures were estimated based on the sequence of listing contained here. When the sequence changes, reductions from each measure could be affected, but the net total remain the same. The purpose of this table is to estimate total emission reductions without overlapping or double-counting between measures.
(2) Total remaining emissions are slightly different from figures in Draft 2003 AQMP main document because of rounding.
(3) Mobility Adjustments include transportation control measures TCM-1A, TCM-1B, TCM-1C and adjustments are reflected in the CEPA baseline beyond year 2000.

ATTACHMENT 5

CEPA Source Level Emissions Reduction Summary for 2010: Annual Average Inventory

Year 2010 Emission Reductions Excluding Natural Sources by Control Measure in the South Coast Air Basin
(Annual Average Inventory - Tons/Day)

(A) Reductions Without Overlapping/Double-Counting With Other Control Measures (1)

		(Reductions - Tons/Day)					
Measure	Name	VOC	NOx	CO	SOx	PM10	PM2.5
BA-2202	Baseline adjustment for R2202	1.83	1.90	20.72	0.02	0.07	0.07
BA-1122	Baseline adjustment for R1122	3.51	0.00	0.00	0.00	0.00	0.00
BA-1113	Baseline adjustment for R1113B	0.00	0.00	0.00	0.00	0.00	0.00
BA-1133	Baseline adjustment for Composting Opr.	-3.68	0.00	0.00	0.00	0.00	0.00
BA-FVR	ARB-Petroleum Marketing - Bas. Adj.	-1.20	0.00	0.00	0.00	0.00	0.00
BA-ARPT	ARB-Airport Ground Support Equip - Bas. Adj.	-0.14	-0.68	0.00	0.00	0.00	0.00
BA-SMOG	ARB-Smog Check II - Bas. Adj.	1.76	3.56	0.00	0.00	0.00	0.00
CTS-07	FER Architectural Ctg (R1113 Phase 3) (VOC)	7.22	0.00	0.00	0.00	0.00	0.00
CTS-10	Misc. Industrial Coating&Solvent Opr. (VOC)	2.79	0.00	0.00	0.00	0.00	0.00
FUG-05	ER from fugitive Emission Sources (VOC)	1.99	0.00	0.00	0.00	0.00	0.00
CMB-09	Petroleum FCCU (PM10,PM2.5,NH3)	0.00	0.00	0.00	0.00	0.27	0.24
BCM-7	FER from Fugitive Dust Sources (PM10)	0.00	0.00	0.00	0.00	0.00	0.00
BCM-08	Aggregate Operation (PM10)	0.00	0.00	0.00	0.00	0.67	0.36
PRC-03	COE fr Restaurant Operations (PM10)	0.00	0.00	0.00	0.00	0.99	0.98
PRC-07	Industrial Process Operations (VOC)	1.74	0.00	0.00	0.00	0.00	0.00
WST-01	Ems Reduction fr Livestock Waste (VOC,NH3)	4.81	0.00	0.00	0.00	0.00	0.00
WST-02	COE fr Composting (PR1133) (VOC,NH3,PM10)	1.20	0.00	0.00	0.00	0.00	0.00
CMB-07	Refinery Flares (VOC,NOX,SOX,CO,PM)	0.16	0.17	0.99	2.16	0.03	0.03
LT1-D	Long-Term Measure 1 : District (2)	10.23	0.00	0.00	0.00	0.00	0.00
MSC-05	Truckstop Electrification (ALL)	0.09	1.69	0.55	0.00	0.02	0.02
CMB-10	Further RECLAIM Reductions	0.00	2.85	0.00	0.00	0.00	0.00
CONS-1	ARB-Consumer Products Limits for 2006	2.27	0.00	0.00	0.00	0.00	0.00
CONS-2	ARB-Consumer Products Limits to 2010	14.69	0.00	0.00	0.00	0.00	0.00
FVR-1	ARB-Vapor from Aboveground Storage Tanks	0.10	0.00	0.00	0.00	0.00	0.00
FVR-2	ARB-Vapor from Gasoline Dispensing at Marinas	0.10	0.00	0.00	0.00	0.00	0.00
FVR-3	ARB-Gasoline Dispenser Hoses	0.70	0.00	0.00	0.00	0.00	0.00
LMD-1	ARB-Passenger Cars + Light Duty Trucks	21.52	14.84	0.00	0.00	0.00	0.00
LMD-2	ARB-Smog Check Improvements	5.64	8.82	0.00	0.00	0.00	0.00
ONHD-1	ARB-Truck and Bus Highway Inspections	0.12	0.00	0.00	0.00	0.11	0.09
ONHD-2	ARB-Vapor from Gasoline Cargo Tanks	1.57	0.00	0.00	0.00	0.00	0.00
ONHD-3	ARB-Clean-up Existing Truck/Bus Fleet (3)	5.08	13.25	0.00	0.00	2.02	1.64
OFCI-1	ARB-New Off-Road CI Standards (Diesel)	0.29	1.25	0.00	0.00	0.47	0.43
OFCI-2	ARB-Clean-up Existing CI Engines (Diesel)	7.49	0.00	0.00	0.00	4.83	4.44
OFCI-3	ARB-Off-Road Equipment Inspection Program (Diesel)	0.00	0.00	0.00	0.00	0.00	0.00
OFLSI-1	ARB-Off-Road New Standards (Gasoline + Nat. Gas)	0.00	0.80	6.75	0.00	0.00	0.00
OFLSI-2	ARB-Clean-up Existing Off-Road Equip.(Gasoline+Nat.Gas)	1.51	3.25	0.00	0.00	0.00	0.00
OFLSI-3	ARB-Electrified New Forklifts (Gasoline + Nat. Gas)	1.03	4.13	8.08	0.00	0.03	0.03
SMOF-1	ARB-Handheld Lawn & Garden Equipment Standards	0.66	0.07	0.00	0.00	0.00	0.00
SMOF-2	ARB-Non-Handheld Lawn & Garden Equipment Stds.	4.52	0.55	0.00	0.00	0.00	0.00
MARINE-1	EPA-Harbor Craft and Ocean-Going Ship Stds. (3)	0.40	3.06	0.00	0.00	0.30	0.28
MARINE-2	ARB-Clean-up Existing Harbor Craft	0.09	2.58	0.00	0.00	0.05	0.04
MARINE-3	EPA-Clean-up Existing Ocean-Going Ships (3)	1.47	16.61	0.00	0.00	1.21	1.12
ARPT-1	EPA-Reductions from Jet Aircraft (3)	0.56	1.80	0.00	0.00	0.00	0.00
LT1-ND	Long-Term Measure 1 : Non-District (2)	40.98	136.58	0.00	0.00	0.00	0.00
LT2	Long Term Measure 2 (2)	158.84	0.00	0.00	0.00	0.00	0.00
Grand Total (Net)		301.93	217.06	37.09	2.18	11.07	9.77

EMISSION SUMMARY FOR
(POINT, AREA, MOBILE SOURCE, AND OFF-ROAD MV)

BASELINE EMISSIONS	VOC	NOx	CO	SOx	PM10	PM2.5
Point source	70.26	12.22	44.18	6.04	13.13	11.70
Area source	229.52	44.42	175.55	1.00	249.96	72.22
RECLAIM	0.00	34.20	0.00	12.03	0.00	0.00
Total Stationary	299.78	90.84	219.73	19.07	263.10	83.92
On-road	195.13	419.10	1981.36	1.94	18.12	11.83
Off-road	91.30	246.23	773.09	37.19	17.95	15.86
TOTAL	586.21	756.18	2974.18	58.20	299.17	111.61
EMISSION REDUCTIONS						
Point source	18.70	3.02	0.99	2.16	0.60	0.45
Area source	107.01	0.00	0.00	0.00	1.36	1.16
Total Stationary	125.70	3.02	0.99	2.16	1.96	1.61
On-road	125.15	129.54	20.72	0.02	2.20	1.79
Off-road	51.08	84.50	15.38	0.00	6.91	6.36
TOTAL	301.93	217.06	37.09	2.18	11.07	9.77
REMAINING EMISSIONS						
Point source	51.57	12.05	43.19	3.88	12.53	11.25
Area source	122.51	44.42	175.55	1.00	248.61	71.05
RECLAIM	0.00	31.35	0.00	12.03	0.00	0.00
Total Stationary	174.08	87.83	218.74	16.91	261.14	82.31
On-road	69.98	289.56	1960.64	1.93	15.92	10.04
Off-road	40.22	161.75	757.71	37.19	11.04	9.50
TOTAL	284.28	539.12	2937.09	56.01	288.10	101.84
NSR/Set-Aside	5.00	3.00	0.50	1.00	0.85	0.85
GRAND TOTAL (T/D) (4)	289.28	542.12	2937.59	57.01	288.95	102.69
Mobility Adjustments (5)	3.60	0.34	65.67	0.06	0.60	0.46

(1) Emission reductions for individual measures were estimated based on the sequence of listing contained here. When the sequence changes, reductions from each measure could be affected, but the net total remain the same. The purpose of this table is to estimate total emission reductions without overlapping or double-counting between measures.
(2) Emission reductions associated with long-term measures, presented under three control measures here, will be updated in the Final 2003 AQMP.
(3) Emission reductions for these control measures affecting federal sources are also considered under long-term strategy. Control measure ONHD-3 targets heavy-duty trucks under state and federal jurisdiction.
(4) Total remaining emissions are slightly different from figures in Draft 2003 AQMP main document because of rounding.
(5) Mobility Adjustments include transportation control measures TCM-1A, TCM-1B, TCM-1C and adjustments are reflected in the CEPA baseline beyond year 2000.

ATTACHMENT 6

CEPA Source Level Emissions Reduction Summary for 2006: Planning Inventory

Year 2010 Emission Reductions Excluding Natural Sources by Control Measure in the South Coast Air Basin
(Planning Inventory - Tons/Day)

(A) Reductions Without Overlapping/Double-Counting With Other Control Measures (1)

		(Reductions - Tons/Day)			
Measure	Name	VOC	NOx	CO	NO2
BA-2202	Baseline adjustment for R2202	1.83	1.78	17.52	2.06
BA-1122	Baseline adjustment for R1122	3.52	0.00	0.00	0.00
BA-1113	Baseline adjustment for R1113B	0.00	0.00	0.00	0.00
BA-1133	Baseline adjustment for Composting Opr.	-3.68	0.00	0.00	0.00
BA-FVR	ARB-Petroleum Marketing - Bas. Adj.	-1.20	0.00	0.00	0.00
BA-ARPT	ARB-Airport Ground Support Equip - Bas. Adj.	-0.14	-0.67	0.00	-0.68
BA-SMOG	ARB-Smog Check II - Bas. Adj.	1.87	3.34	0.00	3.87
CTS-07	FER Architectural Ctg (R1113 Phase 3) (VOC)	8.52	0.00	0.00	0.00
CTS-10	Misc. Industrial Coating&Solvent Opr. (VOC)	3.00	0.00	0.00	0.00
FUG-05	ER from fugitive Emission Sources (VOC)	2.01	0.00	0.00	0.00
CMB-09	Petroleum FCCU (PM10,PM2.5,NH3)	0.00	0.00	0.00	0.00
BCM-7	FER from Fugitive Dust Sources (PM10)	0.00	0.00	0.00	0.00
BCM-08	Aggregate Operation (PM10)	0.00	0.00	0.00	0.00
PRC-03	COE fr Restaurant Operations (PM10)	0.00	0.00	0.00	0.00
PRC-07	Industrial Process Operations (VOC)	1.95	0.00	0.00	0.00
WST-01	Ems Reduction fr Livestock Waste (VOC,NH3)	4.81	0.00	0.00	0.00
WST-02	COE fr Composting (PR1133) (VOC,NH3,PM10)	1.20	0.00	0.00	0.00
CMB-07	Refinery Flares (VOC,NOX,SOX,CO,PM)	0.16	0.17	0.99	0.17
LT1-D	Long-Term Measure 1 : District (2)	10.99	0.00	0.00	0.00
MSC-05	Truckstop Electrification (ALL)	0.11	2.07	0.43	1.31
CMB-10	Further RECLAIM Reductions	0.00	3.00	0.00	3.00
CONS-1	ARB-Consumer Products Limits for 2006	2.27	0.00	0.00	0.00
CONS-2	ARB-Consumer Products Limits to 2010	14.69	0.00	0.00	0.00
FVR-1	ARB-Vapor from Aboveground Storage Tanks	0.10	0.00	0.00	0.00
FVR-2	ARB-Vapor from Gasoline Dispensing at Marinas	0.10	0.00	0.00	0.00
FVR-3	ARB-Gasoline Dispenser Hoses	0.70	0.00	0.00	0.00
LMD-1	ARB-Passenger Cars + Light Duty Trucks	22.80	13.93	0.00	16.11
LMD-2	ARB-Smog Check Improvements	5.95	8.26	0.00	9.60
ONHD-1	ARB-Truck and Bus Highway Inspections	0.12	0.00	0.00	0.00
ONHD-2	ARB-Vapor from Gasoline Cargo Tanks	1.57	0.00	0.00	0.00
ONHD-3	ARB-Clean-up Existing Truck/Bus Fleet (3)	5.07	12.76	0.00	13.93
OFCI-1	ARB-New Off-Road CI Standards (Diesel)	0.30	1.27	0.00	1.24
OFCI-2	ARB-Clean-up Existing CI Engines (Diesel)	7.63	0.00	0.00	0.00
OFCI-3	ARB-Off-Road Equipment Inspection Program (Diesel)	0.00	0.00	0.00	0.00
OFLSI-1	ARB-Off-Road New Standards (Gasoline + Nat. Gas)	0.00	0.82	6.26	0.77
OFLSI-2	ARB-Clean-up Existing Off-Road Equip.(Gasoline+Nat.Gas)	1.60	3.36	0.00	3.15
OFLSI-3	ARB-Electrified New Forklifts (Gasoline + Nat. Gas)	1.03	4.09	8.16	4.17
SMOF-1	ARB-Handheld Lawn & Garden Equipment Standards	0.70	0.08	0.00	0.06
SMOF-2	ARB-Non-Handheld Lawn & Garden Equipment Stds.	4.60	0.61	0.00	0.49
MARINE-1	EPA-Harbor Craft and Ocean-Going Ship Stds. (3)	0.40	3.06	0.00	3.06
MARINE-2	ARB-Clean-up Existing Harbor Craft	0.09	2.58	0.00	2.58
MARINE-3	EPA-Clean-up Existing Ocean-Going Ships (3)	1.47	16.61	0.00	16.61
ARPT-1	EPA-Reductions from Jet Aircraft (3)	0.57	1.84	0.00	1.76
LT1-ND	Long-Term Measure 1 : Non-District (2)	45.06	134.43	0.00	140.51
LT2	Long-Term Measure 2 (2)	170.39	0.00	0.00	0.00
Grand Total (Net)		322.13	213.38	33.35	223.75

EMISSION SUMMARY FOR
(POINT, AREA, MOBILE SOURCE, AND OFF-ROAD MV)

BASELINE EMISSIONS

	VOC	NOx	CO	NO2
Point source	83.25	14.40	47.69	14.40
Area source	231.40	35.43	339.14	57.33
RECLAIM	0.00	35.67	0.00	35.67
Total Stationary	314.65	85.51	386.82	107.40
On-road	204.38	398.53	1714.22	447.90
Off-road	106.01	255.83	626.19	237.04
TOTAL	625.05	739.87	2727.23	792.34

EMISSION REDUCTIONS

	VOC	NOx	CO	NO2
Point source	22.22	3.17	0.99	3.17
Area source	109.45	0.00	0.00	0.00
Total Stationary	131.67	3.17	0.99	3.17
On-road	131.10	123.02	17.52	138.67
Off-road	59.36	87.19	14.84	81.92
TOTAL	322.13	213.38	33.35	223.75

REMAINING EMISSIONS

	VOC	NOx	CO	NO2
Point source	61.03	14.24	46.70	14.24
Area source	121.95	35.43	339.14	57.33
RECLAIM	0.00	32.67	0.00	32.67
Total Stationary	182.98	82.34	385.83	104.24
On-road	73.28	275.51	1696.70	309.23
Off-road	46.65	168.64	611.35	155.12
TOTAL	302.92	526.49	2693.87	568.59
NSR/Set-Aside	5.00	3.00	0.50	3.00
GRAND TOTAL (T/D) (4)	307.92	529.49	2694.37	571.59
Mobility Adjustments (5)	3.73	0.21	57.11	0.51

(1) Emission reductions for individual measures were estimated based on the sequence of listing contained here. When the sequence changes, reductions from each measure could be affected, but the net total remain the same. The purpose of this table is to estimate total emission reductions without overlapping or double-counting between measures.
(2) Emission reductions associated with long-term measures, presented under three control measures here, will be updated in the Final 2003 AQMP.
(3) Emission reductions for these control measures affecting federal sources are also considered under long-term strategy. Control measure ONHD-3 targets heavy-duty trucks under state and federal jurisdiction.
(4) Total remaining emissions are slightly different from figures in Draft 2003 AQMP main document because of rounding.
(5) Mobility Adjustments include transportation control measures TCM-1A, TCM-1B, TCM-1C and adjustments are reflected in the CEPA baseline beyond year 2000.

UNIVERSITY OF CALIFORNIA, RIVERSIDE

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

COLLEGE OF ENGINEERING - 022
CENTER FOR ENVIRONMENTAL RESEARCH & TECHNOLOGY
(909) 781-5791 FAX (909) 781-5790

RIVERSIDE, CALIFORNIA 92521-0434

July 2, 2001

John DaMassa
Manager, Control Strategy Modeling Section
California Air Resources Board
P.O. Box 2815
Sacramento, CA 95812

Re: Interagency Agreement No. 98-004, Task Order 7
"Peer Review of ARB Ozone Modeling for Southern California"
William P. L. Carter, Principal Investigator

Dear John:

Enclosed please find a copy of my peer review of the ARB's ozone modeling plan for the 2002 SIP. Electronic versions of the same review have already been sent to you by email. If you did not receive the electronic copies, please let me know as soon as possible. I hope this is useful.

This constitutes the deliverables of the above-referenced task order. As you requested, the University of California, Riverside, Accounting Office will be sending you the invoice for payment under separate cover.

Sincerely,

William P. L. Carter
Research Chemist

Enclosures

cc: CE-CERT Business Office
UCR Accounting Office

PEER REVIEW OF ARB OZONE MODELING FOR SOUTHERN CALIFORNIA

Report to the
California Air Resources Board
Interagency Agreement No. 98-004, Task Order 7

July 2, 2001

William P. L. Carter
Principal Investigator

College of Engineering
Center for Environmental Research and Technology
University of California
Riverside, California 92521

Summary

The California Air Resources Board (CARB)'s air quality modeling procedures that are underway or being proposed for the 2003 ozone SIP are critically reviewed, and areas of concern and recommendations are summarized. This review is based on a reading of the September 18, 2000 draft of the CARB's ozone SIP modeling protocol document, discussions and meetings with the CARB staff in June 2001, and the reviewer's experience in chemical mechanism development and VOC reactivity modeling. This document gives a summary of what this reviewer sees as an ideal ozone SIP modeling procedure given the current state of knowledge and data availability, and discusses general and specific aspects of the current CARB modeling plan in light of these considerations. It is concluded that for the most part the models and modeling procedures being proposed for use represent the state of the art and incorporate significant improvements over past SIP modeling, but there are potentially significant concerns and recommendations. The major recommendations concern using more episodes in the control strategy modeling to represent the distribution of relevant conditions, discontinuing use of model components, such as the Carbon Bond mechanism, that are out-of-date and have known errors and biases, more comprehensive analysis of uncertainty and bias, and improved documentation of the process and results, especially to policymakers. Recommendations are also made concerning the process for review and external input.

Background

In 2003 California is required to submit a State Implementation Plan (SIP) indicating how it intends to meet ambient air quality standards for ozone in Southern California. Air quality modeling, which is used to estimate the extent to which planned emissions changes will result in air quality improvements, is an important component of this SIP. Ozone is not emitted directly, but is formed from a complex series of reactions of oxides of nitrogen (NO_x) in the presence volatile organic compound (VOC) and sunlight. Because of the complexity and nonlinearity of these processes, modeling provides the only reasonably credible means to estimate the effects of proposed emissions changes on ozone, and thus is the only available means estimate the likelihood that planned controls will achieve the ambient ozone standard in the required time period. However, air quality modeling is an evolving science, and current models have many uncertainties that affect the credibility of their predictions. For this reason, it is essential that the modeling procedures used to support the SIP and other regulatory planning ensure not only that the modeling reflects the best available knowledge and science, but also that the results, and their limitations, are appropriately understood when they are used in formulating public policy.

The California Air Resources Board (CARB), which is primarily responsible for technical support for the modeling needed for the ozone SIP, recognizes the importance of their modeling efforts reflecting the best available science. The CARB has been in the forefront of supporting applied and even some basic research needed to improve its ozone modeling. The CARB was a major sponsor of the 1997 Southern California Ozone Study (SCOS97) to collect data needed to support and evaluate ozone modeling in Southern California. The CARB staff has been working with regional stakeholder agencies in developing a modeling program to support their planning needs. Industry groups such as the Coordinating Research Council and the Electric Power Research Institute provided support to SCOS97 and hopefully will be providing input into the modeling process as it proceeds. This input from stakeholder groups is important to the credibility of the regulatory modeling process but by itself is not sufficient because by definition stakeholder groups have an interest in the results that may affect perceptions of objectivity.

External peer review by appropriate experts provides another means of enhancing quality and credibility of the process. Ideally, the reviewer should not be represent a stakeholder in the process and be able to provide input from an objective perspective. As such, this provides the best complement to stakeholder input. In practice, however, most qualified experts would have some stake in the process as well, particularly with regard to research on modeling analyses or developing model components or inputs. This can be addressed in part by using reviewers with differing areas of expertise and by taking the reviewers interests into account when assessing his or her input. In any case, the CARB requested that the University of California (UC) conduct a scientific peer review of their air quality modeling work in support of 2003 clean air planning requirements. The three reviewers chosen were Dr. Richard Turco of UCLA Dr. Robert Harley of UC Berkley, and Dr. William Carter of UC Riverside, the author of this report. Although Dr. Turco was unable to conduct the review because of contracting problems, both Dr. Harley and this reviewer were able to participate. This report gives the results of the review by the author. His background and experience, and potential areas of conflict of interest that must be borne in mind when assessing this review, are discussed in the following section. This is followed by a brief summary of the review process and then the review itself.

Reviewer Background

Qualifications and Experience

My research concerns the gas-phase atmospheric reactions of volatile organic compounds and the assessment of reactivities of VOCs in the atmosphere. This includes developing chemical mechanisms for airshed models, testing and refining these mechanisms using environmental chamber data, utilizing these mechanisms in airshed models to develop ozone reactivity scales for VOCs, and directing environmental chamber programs to provide data to test the mechanisms and evaluate VOC reactivities. Current projects include developing a new environmental chamber facility for more comprehensive evaluation of mechanisms for gas-phase and particle formation reactions, and developing improved experimental methods for VOC reactivity assessment that can be applied to compounds where environmental chamber methods are not suitable. I am the developer of the SAPRC-99 gas-phase chemical mechanism that is considered by some to represent the current state of the art in this area, and developed the Maximum Incremental Reactivity (MIR) ozone reactivity scale that is incorporated in several VOC regulations adopted by the CARB. More information about my research, list of publications and downloadable versions of most of his recent reports and presentations is available at http://www.cert.ucr.edu/~carter.

I am a past member of the CARB's Modeling Advisory Committee and the South Coast Air Quality Management District's Science Advisory Council and is a current member of the Texas Air Research Center Advisory Board. I am also a past peer reviewer of the EPA's RADM2 model development effort, and in that capacity assisted in the evaluation of the RADM2 chemical mechanism. I am a currently an active participant in the Reactivity Research Advisory Group (RRWG), where I served as the leader of the team developing the Reactivity Science Assessment document and research plan. The RRWG is a partnership of industry and regulatory groups formed to coordinate VOC-reactivity relevant research to inform the development of the EPA's VOC regulation policies.

Potential Conflicts of Interest

Most qualified reviewers in areas of air quality modeling derive some benefit from work related to modeling or model development that may be relevant to the modeling efforts being reviewed. As discussed below, this reviewer is no exception in this regard. This needs to be borne in mind when assessing the comments and recommendations made in this review.

A non-negligible portion of my research has been funded by the CARB, and the CARB has been or is planning to use the results of some of this research in its modeling and regulations. Most relevant to this review is the fact that my SAPRC-99 chemical mechanism is one of the two proposed for use in the 2003 SIP ozone modeling. This mechanism was developed under CARB funding, and a solicited proposal I submitted to evaluate and update it for low NO_x conditions is now being considered by the CARB Research Screening Committee. Past CARB projects supported the development and recent updates to the MIR scale for VOC reactivity, the CARB is currently supporting my project to develop new procedures to evaluate VOC reactivity, and CARB support for additional research concerning architectural coatings reactivity is pending. The CARB, along with various private sector groups, has supported a number of environmental chamber projects in my laboratory, though most of my current funding for environmental chamber studies is from the EPA project to develop the new environmental chamber facility for improved mechanism evaluation. However, a limited amount of support for experiments in this facility is included in the pending CARB project to evaluate architectural coatings reactivity, and additional proposals to the CARB for research using this facility are anticipated. Assessments of uncertainties in the ozone SIP modeling that is discussed in this review may well affect the priorities set by the CARB concerning funding this mechanism development, VOC reactivity, and environmental chamber work.

One of the recommendations in this review is to use consistent speciation assignments in model simulations using different chemical mechanisms. I have submitted pre-proposals to the RRWG and the American Chemistry Council to carry out work to address this need, and if this is funded I plan to ask the CARB to collaborate and provide in-kind support for this effort. This was brought up in the peer review meeting discussed in the following section and in follow-up communications with the Planning and Technical Support and Research Division staff.

Another recommendation in this review is to use the RACM chemical mechanism be used at least for diagnostic purposes in its SIP modeling. If this is adopted, the CARB may ask me to evaluate this mechanism against chamber data or to assist in making emissions assignments utilizing existing procedures and databases developed for SAPRC-99.

Review Process

John DaMassa of the CARB staff provided me with the September 18, 2000 draft of the document entitled “Modeling Protocol for Regional 1-Hour and 8-Hour Ozone Modeling in Southern California for the 2003 State Implementation Plans (Draft #3)”. I read the document and provided a series of written questions and possible areas of concerns that can serve as a basis for subsequent discussions to Don McNerny and John DaMassa by email. Brief discussions were held with Arthur Winer of UCLA about the biogenic emissions work and somewhat more extensive discussions were held with Joe Norbeck of UCR about his concerns with the mobile source emissions inventory. I then traveled to Sacramento and spent most of a day in meetings primarily with John DaMassa, Don McNerny, Bruce Jackson, Luis Woodhouse, Paul Allen of the Planning and Technical Support Division, and Dongmin Luo of the Research Division. Eileen McCauley of the Research Division, Ed Yotter from our Mobile Source Analysis Branch, Michael Benjamin from the Emissions Inventory Branch and Jinyou Liang of my staff was also present. Of the Planning and Technical Support Division also attended for part of the time. The matters discussed included the details on the work on the biogenic and mobile source emissions inventory, other inventory issues, the choice of modeling episodes and the factors that need to be considered, meteorological models (primarily to educate the reviewer in this area), the various models proposed to be used (with emphasis on the chemical mechanism), and current problems of model evaluations. This review document was prepared the week following the meeting.

Recommended SIP Modeling Procedure

A summary of what I see as an ideal ozone SIP modeling procedure given the current state of knowledge and data availability is given in this section, along with my overall recommendations both for the current 2003 SIP effort and for longer term goals for future SIPs. My assessments and recommendations for the individual components of the current SIP are summarized in the section following this.

Choice of Modeling Episodes

Ideally, episodes being modeled for SIP and most other planning applications should represent the full distribution of meteorological conditions that are relevant to the problem being assessed. In the case of the ozone SIP these are not only conditions where the highest ozone occurs, but also other conditions, which may have quite different transport patterns, and thus possibly different dependences of emissions on air quality in different areas. The distribution of meteorological conditions is particularly important when considering transport from one region to the other, which is an important part of the overall SIP development process. The episode selection process should be made based on an assessment by meteorologists of the different types of relevant meteorological conditions that occur in Southern California, their frequencies of occurrence, and the types of ozone exceedences that occur in each. This information can then be used to assess an appropriate distribution of modeling episodes that adequately represents this distribution.

The use of an appropriate distribution of modeling episodes is important because air quality in different episodes may respond differently to proposed emissions changes. For example, one of the important factors that needs to be assessed in developing an ozone SIP concerns the relative benefits or dis-benefits of NO_x controls. It is known that different regions in the same episode respond quite differently to NO_x controls, and by the same reasoning it is quite possible that the same region in different episodes may also respond differently. Use of varying episodes is also necessary to test the extent to which the “relative reduction factor” (RRF) varies with conditions, which policy makers need to know if they are relying on this approximation in their planning. Use of different episodes is also necessary for determining the extent to which sensitivities to uncertainties vary with conditions, and for other diagnostic purposes.

The current procedure for choosing modeling episodes for SIP planning is based primarily on considerations of data availability and the magnitude of the peak ozone concentration. With regard to the latter, if only one episode can be modeled, as has been the case in the past, modeling the “worst case” scenario is probably the best compromise. However, a control strategy that meets the standard in the worst-case episode may not necessarily meet the standard in other episodes where present ozone may not be so high. For example, NO_x controls may reduce ozone more in high photochemical reactivity scenarios, where O_3 tends to be NO_x limited, while it may reduce it less or even make it worse in lower photochemical reactivity scenarios where the fact that NO_x slows down the rate of ozone formation may be relatively more important. Only by modeling episodes that represent different current (and potential future) non-attainment conditions can one assess the effectiveness of controlling for the worst-case conditions in improving air quality under more moderate conditions that may occur more frequently.

Data availability is currently the most important factor governing episode selection in the current modeling approach. Episodes with good data availability and quality obviously are the best for model performance evaluation, but they may not be the best or sufficient in number to represent all conditions that should be considered for control strategy modeling. Just because a certain type of condition is not well characterized for model performance evaluation does not mean that it should be ignored in planning applications. Representing such conditions approximately is better than ignoring them entirely.

One way to address this problem is to use the episodes with good data availability for model evaluation, but supplement them with other episodes developed using generally available meteorological information and expertise to represent the other meteorological conditions that need to be included in a comprehensive planning analysis. The episodes with good input and evaluation data can be used to develop and test the model inputs and components that are common to all episodes, which would include the emissions inputs and the ability to predict their chemical transformations. These are probably the most important uncertainties in predicting relative reduction factors caused by emissions changes. The main model components that change when modeling different episodes are meteorology and to a lesser extent boundary conditions. Uncertainties in boundary conditions can be minimized by appropriately expanding the domain, the approach adopted by the ARB in the present modeling plan. Although uncertainties in meteorological representation are critically important in model evaluations involving comparing predictions to ambient measurement data, they are less important in planning modeling where the main objective is to predict how emissions changes will affect future air quality in general *types* of episodes. Modeling past episodes for evaluation is the means to that end, but not the end in itself.

Use of a distribution of variety of approximately represented conditions has to be used for predicting long term averages, which though not applicable to the present ozone SIP is applicable to other planning applications. It is also the approach used when developing the MIR and my other VOC reactivity scales that are intended for application under a range of environmental conditions. Research may be needed to improve our confidence in developing reasonable, consistent, and credible meteorological inputs representing conditions with only routine supporting data, and without excessive cost in terms of person-hours and computer time. I am not sufficiently knowledgeable in meteorology to assess this, but prognostic modeling appears to me to be the appropriate approach. If this is not currently feasible, research in this area needs to be a priority for future ozone SIPs, as well as for the other planning modeling applications as indicated above.

The draft modeling protocol lists six episodes being considered, with all but one being based on data obtained in the SCOS97 campaign. The CARB staff believes that this represents a good distribution of "episode types" based on their classifications of episodes in terms of transport between the various urban areas in Southern California. This is a significant improvement over most previous SIP modeling in that at least more than one episode is being used, and that an attempt is being made to represent different episode. However, the document does not contain a discussion or a reference to previous work for an analysis of episode types, and whether other considerations besides inter-urban transport might affect how air quality depends on emissions. Certain episodes such as September 26-29 and October 30-November 1, 1997 appear to be de-emphasized because their ozone concentrations are not as high in the SCAB, but they appear to me to be sufficiently different from the others that they should be in a comprehensive analysis. These episodes will be modeled as part of other projects being undertaken by the CARB in conjunction with the regional agencies, so they might as well be included in the ozone SIP modeling. At a minimum they should be used to evaluate the applicability of RRF assumptions.

Treatment of Uncertainty

Uncertainty is unavoidable in model applications, but if their effects on model predictions can be quantified, then planners can take them appropriately into account when using the model results when recommending public policy. It is critical that the modeling plan include an attempt to assess the effects of the magnitudes of the important known uncertainties, and that the documentation on the results of the modeling analysis explain these uncertainties in a manner that is understandable and useable to the policymakers and their advisors that will be using the results. The CARB clearly understands the importance of this, as evidenced by the fact that the modeling plan includes an extensive set of diagnostic simulations and that the CARB has funded work on quantifying uncertainties. However, additional work and better documentation in this area may be appropriate as part of the SIP modeling process.

Quantifying uncertainty is difficult, but it is not infeasible for at least some of the major model inputs and components that may be important in affecting policy-relevant results, particularly those, such as mass emissions, based on scalar values. At least subjective but expert numerical estimates of uncertainty can and should be made for important inputs such as emissions, boundary conditions, light intensity, the more important uncertain parameters used in the chemical mechanism and probably other types of parameters or inputs in the model. This will allow for systematic assessments of the effects of these uncertainties on model predictions. Because of their potential importance, such estimates should be ideally included in the modeling input databases. Contractors developing emissions databases should be required to provide uncertainty ranges as part of the overall projects, and results of mechanism uncertainty studies such as by Milford and co-workers could be used to determine the most important *known* quantifiable mechanism uncertainties could be used. Box models or other types of appropriate simplified models or analysis methods could be used to assess relative importances of the various types of uncertainties to determine which subset or combination of parameter values would be the most useful to examine in diagnostic calculations using the comprehensive models.

This type of systematic analysis is probably not feasible for meteorological inputs or other model components that are multi-dimensional, or for model components where it is uncertain whether the parameterization is appropriate, such as aspects of the chemical mechanism or some of the numerical algorithms. In this case diagnostic simulations using a reasonably varied set of inputs, assumptions, or methods are needed. Ideally, this should be done with only one type of uncertainty varied at a time. This way one gets a more unambiguous assessment of which type of uncertainty is most important and the contribution of that type of uncertainty to the overall result, and there is less chance of errors of one type compensating for errors in another. This will also give a better idea of priorities for future model improvements. For example, comparing simulations with entirely different models, with different chemical mechanisms, resolutions, solvers and emissions processing assignments is much less useful than using simulations where, for example, only the mechanism or only the resolutions or only the solvers are changed. Considerations with regard to particular types of uncertainty are discussed where appropriate in the following section. In many cases limitation of current model software inhibits systematic assessments of different types of uncertainties, and software improvements to make such assessments more feasible are needed. This is also discussed later.

Treatment of meteorological uncertainty is a special case because considerations of this type of uncertainty are quite different in policy-relevant calculations than in simulations of past episodes carried out for model evaluation. In the latter case meteorological uncertainties are extremely important because they have large effects on predicted concentrations at specific times and locations, and must be properly taken into account when evaluating model performance. A decision on whether non-meteorological model components such as the treatment of emissions or chemistry are or are not consistent with the data must be made based on an assessment of meteorological uncertainty and how it affects the predictions of the observations. It is my understanding that major efforts are made in the current and planned model evaluation process to properly consider the effects of meteorological uncertainty on evaluation results, but I am not fully qualified to evaluate their adequacy. However, specific concerns in this regard are discussed later in this report.

As discussed above, meteorological uncertainty concerning exact representations of specific past episodes are much less important in planning modeling, except to the extent that the evaluation results affects conclusions concerning the uncertainties of the non-meteorological model inputs and components. For planning modeling, the most relevant meteorological uncertainties concern how well the selected episode(s) reflect the distribution of conditions of relevance to the planning assessment, and the variability of policy-relevant results with respect to meteorological conditions. Simulations of episodes representing differing meteorological conditions provide the means for assessing this, as discussed in the previous section. This is an important part of the planning modeling procedure that has not been adequately utilized in the past. As indicated above, the current SIP modeling plan represents a significant improvement in this regard, but its practice of only using the most data-intensive episodes for planning modeling means that the representativeness of the episodes used is more a matter of serendipity than planning.

Treatment of Biases

Bias refers to a type of uncertainty where the input or component is considered to be significantly more likely to err in one direction than the other. The now-classic example concerns past vehicle emissions inventories, which many if not most experts thought were low, and few if any thought were high. This may not be the case for vehicle emissions inventories now (some think the bias may be in the other direction), but may be the case in other aspects of the model inputs or components. Examples include treatment of unknowns in biogenic and anthropogenic inventories, known or probable biases in the Carbon Bond mechanism proposed for use in some of the modeling, possible inventory biases that might be suggested by comparing model calculations with measurements of emitted species, and probably others that I am not currently aware of. These specific concerns are discussed where appropriate below. In this section, I give comments that I think should be applicable to all types of biases.

Ideally, known biases in the emissions inventories and where feasible other areas should be eliminated from the models before they are used for planning. This can be done by adjusting the biased inputs until approximately an equal number of experts think they are too high as too low. Procedures for doing this, probably by using advisory panels, should be included in the modeling protocol. Diagnostic calculations and uncertainty analyses could then be done to assess the magnitudes of the corrections and uncertainties. Removing biases in this way may increase the overall modeling uncertainty, but uncertain models without known biases are better for policy analyses than biased ones. The documentation of the results should include a discussion of the adjustments that had to be made, and the how the adjustments and uncertainties affect the policy-relevant results. If the biases are large and the adjustments to remove them affect policy-relevant conclusions, this is something that the policy makers definitely need to understand.

The contractors responsible for developing components of the emissions inventories should obviously be among the experts used when assessing inventory biases, though they probably should not be the only ones. As indicated above, such contractors should be required to include estimates of uncertainties in their data. As part of that, they should also be required to indicate where they think biases may exist and to provide estimates as to the likely magnitudes of the biases. Furthermore, as discussed below, they should not be allowed to treat unknowns considered to be non-negligible as if they do not exist. They should provide their best estimates as to the amount of unknowns present and their uncertainty range, and make recommendations on how they should be represented in the model (i.e., what types of chemical species might be appropriate surrogates for them). Lack of knowledge should not be accepted as an excuse for not providing such recommendations. Proper modeling procedure requires that unknowns that may have non-negligible effects be represented somehow, and the contractors familiar with how the data are obtained usually can make more educated (or less uneducated) guesses about the data than the modeler.

Some types of known or potential biases may not be possible to assess using the types of procedures discussed above. Model components with known biases should not be used in planning calculations if better alternatives are available, though in some cases simulations with the biased component may be useful for diagnostic purposes. (An example is the chemical mechanism, which is discussed below.) Whenever possible, effects of biases that cannot be readily removed should be examined with appropriate diagnostic calculations to determine the types of effects they may have on results of planning calculations. Based on this, the likely biases in terms of policy-relevant predictions should be explained in the documentation so they are understood by those using the results. If it is not possible to assess the biases in this way, the documentation should still indicate the existence of the biases, and the direction and reasonable upper limit magnitude for the resulting bias in the policy-relevant result. Known biases should never be treated as if they do not exist.

State of Science of Model Components

The CARB recognizes the importance of its planning models representing the state of the science, as indicated by past research projects and its planned use of the most advanced models, so this need not be discussed in detail here. Whenever feasible it is best to use more than one approach and representation for an uncertain model components such as the chemical mechanism, treatment of the meteorology, and numerical approximations used in the model software, so effects of uncertainties that are otherwise difficult to quantify can be examined. Ideally, the alternative approaches should be as different as possible yet equally represent the state of the art, though this ideal is difficult to achieve in most cases. However, the more common approach is to compare a state-of-the-art model or component with one that is known to be outdated. This is better than no comparison at all, but does not necessarily indicate the uncertainty of current model, only the extent to which advances in our knowledge since the latter was developed have changed model predictions. Such comparisons may give misleading uncertainty assessments without a detailed consideration of the nature of the updates and the remaining uncertainties. Discussions of specific components that I am aware of are given below.

The CARB's modeling plans include use of different models and in some cases different model components in their performance evaluation studies. Some of the models and components discussed in the modeling protocol document reflect more of the state of the art than the others. The less advanced models should probably not be used for planning purposes, though will probably be useful for diagnostic purposes. Although as discussed above it is better that errors in model components be assessed systematically by comparing state-of-the-art alternatives and by varying one component at a time, the CARB's planned comparisons are better than nothing when such a systematic evaluation is not feasible, and may provide the most straightforward way to test for gross problems and errors.

An important model component that may not be receiving adequate attention is the model software. The internal software design and code documentation should be such that programming improvements can be carried out relatively easily by programmers who did not work on the original code.

Just as the CARB should not use proprietary models for regulatory applications, so should it not use models whose code cannot be understood or appropriately by qualified CARB staff or independent reviewers and contractors. All contractors doing software development work for the CARB should be required to produce well documented and publicly available source-code, and software developed internally by CARB staff should be held to the same standards.

Software limitations that prevent useful diagnostic simulations from being carried out or that introduce known errors into the model should be removed to the extent feasible. For example, most models and emissions processing software systems do not readily permit use of different chemical mechanisms in a consistent manner, and errors are introduced into model simulations because the software does not permit input of aircraft emissions into elevated layers. The chemical mechanism implementation modules should be such that new chemical mechanisms, including those using variable parameters that depend on the compounds being represented, can be readily implemented and the implementation readily verified. The emissions processing software and database design must allow efficient processing of emissions for different mechanisms in a consistent manner. This will not only aid in the implementation of updated mechanisms in those models where this is needed, it will also improved capabilities for diagnostic calculations to evaluate chemical mechanism uncertainty.

Model Performance Evaluation

Because of the many uncertainties involved in the model application, model performance evaluation is necessary, though not sufficient, to provide some degree of credibility to its predictions. The theory is that if a state-of-the-science model with credible inputs can correctly simulate observations in a past episode, then it *may* be able to simulate what happens if emissions change in some future episode. The CARB recognizes the critical importance of model evaluation, so this need not be discussed in further detail here. The CARB also recognizes that there is always concern for compensating errors or "giving the right answer for the wrong reasons", and that this is why the model must utilize the best science and inputs available. This is also why any adjustments or modifications made to the model specifically to improve its performance must be made with great care, and with an appreciation that the adjustments may be covering up an error of a totally different type than the parameter being adjusted,

Model performance evaluation should be carried out with a proper appreciation of the critical role of meteorological uncertainty in the model evaluation process. This is because predictions of observations at specific sites are highly sensitive to meteorological inputs, but meteorological inputs are held constant when carrying out control strategy predictions. Indeed, meteorological uncertainties are actually of secondary importance in control strategy modeling, since the objective is just to represent a *type* of episode, not an actual past event. Because of this, inappropriate adjustment of meteorological parameters to compensate for errors in the emissions or chemistry may result in a model that is actually *less* accurate in control strategy predictions than might otherwise be the case.

The modeling protocol document does not include adequate discussion of the steps being taken to assure that assure that meteorological input adjustments are not being made to compensate for non-meteorological errors. Adjustments to fit purely meteorological observations such as wind or temperature fields are clearly appropriate if done in a physically reasonable manner, but adjustments to fit observations of pollutants must be done with greater care and always be clearly documented. Adjustments to fit slowly reacting primary species such as CO or morning NO_x might be appropriate if their emissions are reasonably well established and if there is independent evidence supporting the adjustment. Adjustments to fit secondary species such as O_3 should be done with much greater care, and probably should be avoided unless it can clearly justified and supported by other observations.

Diagnostic simulations to determine effects of the adjustments on control strategy predictions should be conducted if meteorological adjustments that may compensate for errors in emissions inventories or chemical mechanisms have to be made to obtain satisfactory model performance.

There is always the possibility that the model performance will be unsatisfactory in some respect, unless adjustments that are otherwise difficult to justify are made. Ideally all the different types of uncertainties that may cause the bias should be examined, to determine the effects of the different types of errors on control strategy predictions. For example, if the model does not correctly predict the morning concentrations of NO_x, the error could be in the emissions or the emissions or in the mixing heights. Control strategy predictions using models with the meteorological or with emissions adjustments can be compared to assess the uncertainty in the policy-relevant result indicated by this model performance problem. If the model predicts that the ozone is too low in mid-basin regions, the error could be in the emissions, the chemical transformation rates (i.e. the mechanism or the light model), or the winds. Each of these kinds of adjustments should be assessed to see if they improve model performance, and if so what differences they cause in policy-relevant predictions. Although the appropriate adjustment may be uncertain, at least policy makers would get some idea of the uncertainty in the model predictions that may be indicated by the poor model performance.

Since the actual objective of the SIP modeling process is to predict effects of emissions changes on air quality, the most directly relevant model performance test wold be to see if the model actually predicts the results of past and ongoing emissions changes. Emissions changes occur every weekend, and it is now recognized that weekends have statistically different air quality than weekdays. Indeed, weekend/weekday ozone ratios have been used as an indicator of the relative sensitive of ozone to NO_x and VOC controls. Therefore, a very useful performance evaluation test would be to see if the model could predict these weekend/weekday differences. This is can be done by modeling weekday episodes using weekend emissions, and vise-versa. Such an assessment is feasible now because the CARB now has separate weekend inventories, and certainly should be carried out for all the episodes being considered. However, at present but the applicability of the results may be too uncertain to draw definitive model performance conclusions because of the uncertainty in the weekend inventories, and because the limited number of model scenarios does not allow for an adequate comparison with statistical air quality data. Reducing the uncertainty of this very useful type of model performance evaluation is one reason that improving the quality of weekend inventories should be a high priority. This is also another reason why a more comprehensive set of model scenarios needs to be developed, as discussed above.

Documentation

Proper documentation is a critical component of the modeling process. It serves two important functions. The first is to convey to the policymakers the policy-relevant results of the analysis and the assumptions, caveats and uncertainties involved. The second is to document the process for technical experts in the stakeholder and research community, and for the public record. These are essentially separate documents, though they should be consistent and lead to the same conclusions.

The executive summary and recommendations for the policymakers is probably the most important output of this process, because it determines what the public gets for all this effort that it paid for. This is also probably the most difficult to produce properly. Not only must the general methods and the results be given in an understandable and meaningful way to a largely non-technical audience, it must make this audience appreciate and as much as possible understand the uncertainties and potential biases involved. If insufficient information or caveats are provided than the results may be misused and inappropriate decisions may be made. If too much detail is given then the information will probably not be read or assimilated and the result would be the same if it had not been provided in the first place.

This reviewer is not an expert in preparing executive summaries for policymakers, and therefore cannot provide specific recommendations other than that a significant effort be expended to see that this is done as well as possible. One suggestion is to use members of the CARB or stakeholder legal staffs as test subjects to serve as surrogates for the intelligent but not necessarily technically trained audience that the summary is designed to reach. The subjects would then be given a limited amount of time to read the summary and then asked to present their assessments of the results and their uncertainties. If the subjects misinterpret or cannot properly assimilate the information in the amount of time that one might reasonably expect a policymaker to spend on it, then the summary would need to be revised and re-evaluated using another set of test subjects. The cost and effort of this procedure may well be worth it in terms of how well the overall modeling process ultimately benefits the public.

The technical documentation and to some extent the policymaker summary should include a discussion of the episode types used in the modeling evaluation in terms of all relevant meteorological variables, and a discussion of the extent to which the episodes used for modeling represents this distribution of episode types. This needs to know this when assessing the implications of the modeling results. All adjustments made to the model to improve model performance, other than error fixes or corrections to input data, need to be listed and clearly documented. There exists (fairly or unfairly) a perception that the models contain so many ad-hoc adjustments to give desired evaluation results that the evaluation is not a real indicator of model validity. One way to combat this is to make the adjustment process completely transparent. Various other items that are recommended for the technical documentation and/or executive summary are indicated where relevant elsewhere in this review. The other components of an appropriate technical documentation are probably adequately known to the CARB staff, and need not be listed here.

Discussion of Specific Components

This section gives my comments and recommendations about specific components of the modeling plan as described in the protocol document and as discussed with the CARB staff, in light of the more general comments given in the previous section. The extent of discussion of the various components reflects primarily my level of familiarity with the components, and not necessarily their relative importances.

Modeling Episodes and Domain

The expansion of the modeling domain to cover the major Southern California source areas represents a major improvement over previous modeling procedures. The domain appears to appropriately minimize problems with boundary conditions, since the most of the boundary areas have relatively low emissions. Inclusion of the urban areas in Northern Mexico is appropriate, although the greater uncertainty of Mexican emissions is uncertain, this approach is treating Mexican pollution as a boundary condition. The inclusion of the other Southern California Air Basins besides the SCAB permits improved planning for the other basins and better assessment of transport issues, as the protocol document indicates.

Southern California appears to be fortunate compared to the East Coast in that it is less important to model half the continent to adequately deal with the major transport and boundary issues. However, the documentation should include some discussion of the extent to which Southern California gets pollution from very long-range transport outside this domain. For example, I understand that PM pollution from Asia can occasionally be significant. Although this may not be an ozone issue, it suggests that very long-range transport may be non-negligible, especially in future episodes if U.S. pollution decreases faster than in other countries.

As discussed above, the choice of episodes for modeling appears to be based more on evaluation issues than on planning considerations. However, the 1997-1998 episodes listed in the protocol document at least represent different transport conditions, and as such provide some ability to test how control strategy effectiveness may vary with meteorological conditions. It may not be feasible to develop entirely new episodes for the current SIP, even if they are used only for planning and rely on modeling of the other episodes for evaluation of the non-meteorological components. However, it may be feasible to us episodes set up previously for modeling to represent other meteorological conditions in the planning modeling for the current SIP. The main difficulty may be expanding the domain to the extent used in the SCOS97 episodes. If it is feasible to obtain credible meteorological inputs (at least in terms of representativeness of a *type* of meteorology) for these past episodes, the SCOS97-evaluated emissions, chemistry, and models can be used to represent these conditions in planning modeling. The feasibility of this should be explored for the current modeling plan. Obviously, this is only worth considering if the meteorological conditions they represent are sufficiently different to complement those in the current episodes.

In any case, as discussed above the SCOS97 episodes considered of secondary importance for SCAB SIP modeling in the discussion in the protocol document should be included as part of the SCAB SIP modeling, if only to provide a more comprehensive test of the RRF approach. Because of the nonlinearity of the processes, the RRF may be quite different in different episodes.

Chemical Mechanism

In terms of predictions of effects of future emissions changes on ozone, the chemical mechanism is probably the second most important model component after the emissions. (Meteorology is probably more important than the mechanism in the performance evaluation modeling, but as indicated above that is a separate process than control strategy modeling.) The CARB recognizes the importance of the chemical mechanism (at least for VOC reactivity modeling) and has been the only significant supporter of chemical mechanism development in the United States in the last decade.

The current modeling plans include use of models with both the Carbon Bond IV (CB4) and the SAPRC-99 mechanisms. Use of SAPRC-99 is appropriate since the CARB supported its development and peer review for VOC reactivity modeling, and has incorporated or is considering incorporating a VOC reactivity scale based on it some of its stationary source regulations. The CARB's Reactivity Science Advisory Panel supported the conclusions of the peer reviewer that it represented the state of the art for reactivity modeling, and also recommended it be used in the CARB's regional modeling. The main problems with this mechanism are its relatively large computational demands and more its complex emissions processing requirements compared to mechanisms used previously in such models. However, advances in computer technology has made the computational demands relatively less important than in the past, and the CARB staff has already implemented its special emissions processing procedures at least for some models.

Use of the CB4 mechanism is included in the plans because it is by far the most widely used in regulatory modeling in the United States, because it has been optimized for computational efficiency, and because emissions processing procedures and databases have been well established for it for many years. For this reason, modelers have a large body of experience and degree of comfort with this mechanism. However, it was developed in the late 80's and there have been numerous advances in our knowledge of atmospheric chemistry and in modeling reactions of VOCs since that time. Although it was evaluated against chamber data, the data used were relatively poorly characterized outdoor chamber experiments, and the treatment of photolysis reactions in the model is inadequately documented and probably not consistent with their treatment during the evaluation. Despite that, the mechanism has undergone essentially no updates since that time, other than updating PAN rate constants in the mid 80's, fixing a numerical instability problem due to an omission in the peroxy + peroxy chemistry in '93, and most recently updating the representation of isoprene. In addition, the mechanism has known errors and biases that could affect control strategy model predictions. Its treatment of reactions of internal olefins (including terpenes that are important in biogenic inventories) ignores their important reactions with ozone, which Paulson and co-workers propose may be an important source of OH radicals at nighttime and other conditions. Its treatment of large molecules is not consistent with the fact that they usually become stronger radical sinks as they become larger, and is not conducive to appropriately modeling secondary organic aerosol formation. Perhaps more significantly in terms of ozone SIP modeling, the limited number of systematic studies carried out thus far suggest that this mechanism predicts ozone formation occurs significantly slower under relatively high NO_x conditions than is the case for more up-to-date mechanisms that have been more comprehensively evaluated against chamber data. This represents a known bias in planning modeling.

For this reason, and the fact that there are alternatives available that are more up-to-date and have fewer known errors and biases, the CB4 should no longer be used in modeling for policy development. However, diagnostic modeling using CB4 is still appropriate and is probably necessary. Implementing a new mechanism into an airshed model is a complex process that is subject to errors, and comparisons against the well-established and (probably) already debugged CB4-based models provide a useful debugging tool. Care must be taken to assure that causes of any differences between models with updated mechanisms can be understood in terms of differences in the mechanisms themselves, and not to implementation or emissions processing errors. The CB4 mechanism also provides a useful method to test the implementation of “flexible mechanism” software into airshed models where CB4 is “hard-wired”. Furthermore, since CB4-based modeling has been used for almost all U.S. regulatory modeling in the past, comparisons of planning predictions of the newer models with CB4-based models provides a link to the past that will probably be useful for some purposes.

Note that when comparing CB4 with other mechanisms care must be taken to be sure that the comparisons are done on an equal light intensity basis. Failure to do this is apparently a problem with some of the CARB’s mechanism comparison and UAM/CB4 vs. UAM-FCM/CB4 tests.

Because of its known biases and errors, comparisons of diagnostic calculations using CB4 with updated mechanisms do not provide an appropriate indication of the effect of chemical mechanism uncertainty. If the differences observed are primarily due to mechanism updates and improvements, then the comparison may over-estimate the effects of mechanism uncertainty. On the other hand, if some highly uncertain process is represented based on the same assumptions in both mechanisms, then the comparison may under-estimate the effect of mechanism uncertainty. This means that such comparisons are not particularly useful for either upper or lower limit uncertainty determination. Although comparing equally state-of-the-art mechanisms also suffers from the problem that both may make similar assumptions about some uncertain processes, at least the results would be useful for lower limit uncertainty determination.

One of the arguments for using CB4 is its computational efficiency, which makes it greatly preferred by many modelers for routine use, especially for applications where a detailed representation of the chemistry is not considered to be needed. If this is important, then the approach should be develop a condensed version of a current mechanism rather than continue to use an out-of-date one with known errors and biases. The position of the CARB staff appears to be that the advantages of improved computational efficiency is not sufficient to justify the cost of work needed to condense an existing mechanism, especially since existing mechanisms tend to be updated from time to time, making condensed versions of them out-of-date. However, it would not be as out-of-date as CB4. Furthermore, an appropriately designed research program examining effects of various condensation and implementation efficiency improvement approaches for molecularly-based mechanisms could probably provide relatively straightforward and inexpensive procedures for updating condensed in the future. I suspect that developing a new condensed current mechanism is given low priority because use of CB4 is available as an option when computational efficiency is important. The priority may be different if the option of using a mechanism with known biases and errors is removed.

If my recommendation to de-emphasize use of the CB4 is adopted, then as presently proposed the CARB's regulatory modeling will be left with only one mechanism. The complexity and uncertainties in the chemical mechanism, and its importance to ozone modeling, makes reliance on only a single mechanism a source of concern. Fortunately, SAPRC-99 is not the only state-of-the-art mechanism that is currently available for use in regional models. The RACM mechanism of Stockwell and co-workers was developed around the same time frame as SAPRC-99, and is widely used in Europe. Because it does not represent the hundreds of types of VOCs that are represented in SAPRC-99 it is not as well suited for VOC reactivity assessment, but it is sufficiently detailed for regional modeling, being comparable to the fixed parameter version of SAPRC-99 in this regard. In addition, it may have a more accurate representation of the peroxy + peroxy reactions that are important under low NO_x conditions, so it would be particularly useful for comparison with SAPRC-99 under low NO_x conditions. Although it has not been as extensively evaluated against chamber data as SAPRC-99, its evaluation is probably sufficient to establish that it does not have major biases. Most of the emissions processing used for SAPRC-99 could be used to derive emissions assignments for RACM with relatively little effort[1].

Because of this, I recommend that the CARB consider implementing the RACM chemical mechanism in at least one of the state-of-the-art models it is proposing to use, and at a minimum conduct mechanism comparison simulations using it for diagnostic purposes. Because of its use in existing models in Europe and its similarity to the fixed parameter version of SAPRC-99 in terms of structure and emissions assignments, I believe that this should feasible for the 2003 SIP. This should be a priority if significant differences between SAPRC-99 and CB4 are found in the model performance that cannot be attributed to the known deficiencies of CB4. However, even if this is not the case, the different way RACM treats low NO_x chemistry compared to SAPRC-99 and CB4 makes it possible that RACM may give different low NO_x predictions than both these mechanisms. Results of past mechanism intercomparisons with RADM2, which has a similar treatment of low NO_x chemistry as RACM, may provide useful guidance in this regard, if any such intercomparisons exist.

Note that any mechanism comparisons, whether with CB4 or state-of-the-art mechanisms, need to be done on a consistent basis in terms of representing different VOCs species in emissions profiles. This is particularly a problem with CB4 where many profiles are assigned directly to model species without specifying the individual contains.

It should be noted that some of the models simulate quite high altitudes that are beyond the likely range of validity of the current mechanisms, particularly for the reactions of the higher VOCs. This may

[1] The reviewer has a conflict of interest in making these recommendations. See "Reviewer Background" section, above.

not be a concern if the chemistry of higher VOCs is not important at higher altitudes and if the most important reactions at higher altitudes are the simpler ones where the temperature and pressure dependences are known and adequately represented. However, it may be appropriate to consider if there may be cases where this could be a concern.

Air Quality Model Selection Process

The CARB is evaluating use of a number of state-of-the-art airshed models for its 2003 ozone SIP modeling. It has not made the final determination on the ones that will primarily be used, except that the out-of-date CB4 will be de-emphasized and non-public domain models will not be used. This is the appropriate procedure. The CARB recognizes that the version selected should represent the state of the science, including use of a current chemical mechanism. Intercomparison of different models is important, though as discussed above it most useful when the different components can be varied independently. Although the modeling protocol document is somewhat vague on the criteria that will be used to select among the recently developed models, the reviewer is reasonably confident that the CARB staff will probably make the appropriate choices.

Since as discussed above out-of-date mechanisms should only be used when appropriate for diagnostic purposes, If SAPRC-99 or an alternative state-of-the-art mechanism cannot be implemented into the model with the funding and time available for this SIP, then that model should not be used. Mechanism implementation is discussed in more detail below.

Mechanism Implementation in Models

The mechanism implementation capability of some of the models being considered for use have been or are being improved as part of the CARB's modeling efforts. The problems with the chemical solver in the original version of the CALGRID model have apparently been fixed, and updated flexible mechanism implementation software is being developed for it. The CARB is funding the implementation of a fixed-parameter version of the SAPRC-99 mechanism into CAMx, to be compatible with the version that is implemented into Models-3. However, if these implementations lack the capability to readily modify the variable product yield parameters in the SAPRC-99 mechanism, it will not permit use of the full capability of this mechanism to represent effects of changes in current or future emissions inventories.

The CARB spent considerable resources to develop the SAPRC-99 mechanism and support the experiments necessary so it could predict ozone and other impacts of the many types of VOCs that are emitted. The variable parameter features of this mechanism permit this chemical detail to be incorporated into airshed models. This is an important feature for assessing the effects of VOC composition changes in emissions, particularly in future year scenarios where emissions compositions may be quite different. Software implementations that do not permit use of this feature means that a less accurate mechanism has to be used in future year simulations than would otherwise be the case. As indicated elsewhere, every effort should be made to remove from the models software limitations that artificially limit the model's accuracy whenever this is feasible. Certainly development of new mechanism implementation systems that lack this capability should not be supported.

The mechanism implementation system used in the UAM-FCM is optimum for use with the SAPRC mechanisms because it fully supports the use of the variable parameter of the mechanism and because it uses the same mechanism compiler as used when the mechanism was developed and evaluated. This means that updated versions of the mechanism can be implemented readily and with significantly reduced likelihood of errors. Unfortunately, the UAM-FCM is being phased out because of limitations of the UAM as a whole, and the CARB apparently has no plans to implement it into more state-of-the-art models. The original version of CALGRID had an earlier version of the SAPRC implementation software that was used as the starting point in the UAM-FCM development, and it could be updated for use with SAPRC-99 with probably relatively modest effort. However, but the CARB staff is apparently developing an entirely new mechanism implementation system for it that uses a different input format and that may lack variable parameter capability. It seems to me that updating the CALGRID mechanism implementation software to the level of UAM-FCM would have been more cost-effective and more beneficial to the CARB in the long run.

The Models-3 software has a flexible mechanism compiler that allows implementation of fixed parameter mechanisms, and files implementing SAPRC-99 into that system are available. Although it lacks variable parameter capability, the model can be run with variable parameter capability by re-compiling the mechanism whenever the parameters are varied. However, my understanding is that some programming is needed so that it can be used in this way without a large level of effort. This capability may be developed for the EPA under separate funding from the American Chemistry Council, but if this does not occur then the development may need to be carried out by the CARB if it wishes to use the full capabilities of SAPRC-99 with Models-3[2].

I understand that the CARB is funding Environ to implement a fixed parameter version of SAPRC-99 into the CAMx model. The implementation should be such that the mechanism can be readily recompiled with different parameter values without having to have a separate contract with Environ. The CARB should be aware that Yosuke Kimura at the Center for Energy and Environmental Resources at the Univ. of Texas at Austin is working on a adapting the UAM-FCM software into CAMx so it can be used with SAPRC-99 in the full parameter mode, though he is apparently doing this with limited resources had is currently experiencing problems. I recommend that the CARB staff most familiar with the UAM-FCM determine the status of Mr. Kimura's efforts, assess its probability of success, and provide whatever assistance may be useful. A subcontract to Environ to provide technical assistance would probably be the most effective approach, and in the long run would probably be the most cost-effective if the CARB plans to extensively use this model[3]. Alternatively, the necessary funding to complete the FCM implementation in CAMx may be obtained from the State of Texas, in which case the CARB should take advantage of the results[4].

I am not familiar with the mechanism implementation system in SARMAP and the amount of work required to implement SAPRC-99 in that model. I am told that it has a flexible mechanism input capability, but I am unsure of the extent to which it supports use of variable parameter mechanisms.

---

[2] The reviewer has a conflict of interest in making this recommendation since he is participating as a subcontractor in a proposal by MCNC to the ACC to do this implementation. See "Reviewer Background" section, above.

[3] The reviewer has a conflict of interest in making this recommendation because CAMx is implemented at CE-CERT and has collaborative programs with Environ using CAMx, and if CAMx had the FCM capability it would significantly enhance our ability to conduct reactivity-relevant research using a state-of-the art model.

[4] The reviewer is engaged in discussions with Dr. Dave Allen of the University of Texas regarding model development work related to this.

Diffusion and Transport Solvers

No discussion is given as to mathematical problems some models have in representing diffusion and transport. I am not an expert in this area, but I understand from discussions with Akula Venkatram that this may be a serious problem. The CARB staff appears to acknowledge that this may be a problem, but apparently does not feel that this is equally a problem with all current models and that much can be done about it at the current time. However, some considerations of the types and magnitudes of any biases this may introduce in model applications or evaluations need to be carried out and discussed in the modeling support documentation. If it is concluded that this is probably a minor problem then documentation or references supporting this conclusion should be given. If it may significantly affect evaluation or application modeling, then appropriate caveats must be given in the documentation, and appropriate priority should be given for research in this area. Known problems should not be ignored even if they cannot be readily solved.

Horizontal and Vertical Grid Resolution

Although I am not an expert in this area, the horizontal and vertical grid resolution used for the more state-of-the-art models being considered appear to represent an improvement over modeling for past SIPs, and probably represent the state of the art. The resolution is less for the UAM models, but its use is being de-emphasized. However, the need to use finite grid resolution may introduce biases in the evaluation or planning modeling, and either previous studies giving information in this regard should be included in the modeling documentation, or diagnostic calculations where resolution is varied should be included as part of the modeling process. Ideally, simulations of selected evaluation and future-case should be carried out with enhanced resolution to determine if the improvements result in significant differences, and if so in what direction. If this isn't feasible (and I understand that enhancing vertical resolution is difficult) then diagnostic calculations can be carried out with the resolution degraded to obtain information on possible biases by extrapolation.

Meteorological Models and Inputs

I am not an expert in meteorology and am not really able to adequately peer review the CARB's performance and plans in this regard. It is unfortunate that none of the current peer reviewers are experts in meteorology, and a meteorologist should be among the reviewers of the evaluation results (see discussion of Technical Oversight and Review, above.) However, the CARB modeling process has extensive input from meteorologists in stakeholder agencies. My general impression based on previous experience, the modeling protocol document, and discussions with the CARB staff has led me to believe that CARB's current meteorological modeling and input preparation efforts represent the state of the art.

The use of both prognostic and diagnostic models as alternative approaches for preparation of meteorological inputs appears to be an appropriate and prudent procedure for model evaluation, since they are entirely different approaches with presumably different strengths and weaknesses. I get the impression that prognostic models are considered to be more state-of-the art, but may not always be as successful as diagnostic models in reproducing observations. Although data assimilation is used with the prognostic models, apparently it is not always sufficient to make the model adequately agree with the data. It is not clear to me what procedures will be used if the two procedures give significantly different model performance evaluation results, and what procedures will be used if the temperature fields and other meteorological observations are not well predicted by the models.

It seems to me that prognostic models would be most useful for developing inputs for episodes to represent meteorological conditions for which data are insufficient for evaluation modeling. In this case, it would be useful to evaluate whether models using meteorological inputs derived from prognostic models tend to give different control strategy predictions than those with inputs derived using diagnostic models that give better fits to the meteorological observations. My suspicion is that although the predictions of observed pollutant concentrations at fixed locations may be different, the RRF predictions would probably be similar. If they are not, it would suggest that control strategy effectiveness may be highly dependent on scenario conditions, making it all the more important that an adequate distribution of conditions be represented in multiple scenarios. The prognostic model inputs may not be properly representing the historical episode being modeled for evaluation purposes, but it may well be that a some episode in 2010 may be better represented than those inputs than inputs that exactly duplicate a 1997 SCOS episode.

One type of meteorological input that in the past has not received adequate attention is the characterization of the solar light intensity and spectral distribution, and how it varies with time. These are critically important inputs that affect model predictions of how rapidly the overall photooxidation processes occur. Improper treatment of this is exacerbated by use of older models, such as the UAM, which treats photolysis rates as if they only depend on the chemical mechanism and the solar zenith angle. State-of-the-art models (and the UAM-FCM) do not have this problem, and require separate input of the actinic fluxes as part of the scenario conditions.

The CARB has come to recognize the importance of light characterization as a model input, and included a number of light characterization studies as part of SCOS97. An analysis by Vuilleumier and co-workers suggest that there are problems with the NO_2 actinometry measurements made during this study, and work carried out at our laboratories indicate that standard “UV radiometers” do not give consistent measurements and should not be used for model input. However, the other light measurements provide useful model input that should be incorporated in the SCOS modeling. Note that Vuilleumier and co-workers and others find that aerosol pollution probably has a non-negligible effect on photolysis rates and that this should be taken into account for comprehensive modeling. At a minimum, diagnostic calculations should be conducted to determine if this might be a significant factor. In addition to potentially affecting evaluation results, future reductions in haze may cause non-negligible changes in photolysis rates that may affect ozone formation. This may not be a large effect but until it is assessed it will be an uncertainty.

Day-Specific Emissions

The CARB is making a major effort to improve day-specific models, and it appears that good progress has been made. Weekend emissions are still more uncertain than weekday, but I get the impression that this is improving. Reducing uncertainty in weekend emissions is important not only for model evaluation of episodes that occur over weekends, but for evaluating the models’ ability to predict weekend/weekday effects, as discussed above. Estimates of the level of uncertainty in the weekend inventory and diagnostic simulations to determine the effects of these uncertainties may be useful.

I understand that an important uncertainty in this regard is weekend emissions on surface streets (freeways seem to be much better characterized in this regard). Because of the utility in testing model predictions of weekend/weekday effects, I recommend that priority be given to reducing this uncertainty. Although I am not an expert in this area, it seems to me that conducting traffic counts on appropriately selected surface streets would provide the type of data needed. I have not been informed of ongoing research in this area, if any.

Chlorine Emissions and Model Representation

If sufficient chlorine or other chlorine containing species that rapidly photolyze to form chlorine are emitted into an air basin, the chlorine atoms they form can contribute significantly to the rate of ozone formation. Chlorine emissions from industrial sources have been shown to be important in affecting ozone formation in Houston, and it has been proposed that non-negligible amounts of photoreactive chlorine-containing species may be formed in heterogeneous reactions of sea-salt aerosol in Los Angeles. If this is so, its emissions and chemistry must be accounted for in the model. The CARB staff informs me that research is underway to assess chlorine emissions in the SCAB, and that steps will be taken to represent it in the models. Note that if it is important in the coastal areas of the SCAB, it may also be important in other coastal areas such as Santa Barbara or San Diego.

Note that incorporating chlorine chemistry into the model require a major expansion of the chemical mechanism and a significant increase in its size. Researchers in Texas are working on adding a chlorine module to the SAPRC-99 mechanism, and also conducting experiments concerning the effects of chlorine on ozone and other measures of air quality. If the CARB decides it needs to incorporate chlorine in its models it should coordinate with and take advantage of the Texas work to avoid duplication of effort and obtain maximum benefit for the available research funding. Work conducted at our laboratories with several chlorine and bromine-containing compounds suggest that we cannot successfully model all the significant atmospheric reactions of halogen-containing species. Therefore, the ability of mechanisms to accurately represent the effects emissions of chlorine or chlorine-containing species will be highly uncertain.

HONO Emissions

HONO may be either emitted or heterogeneously formed from primary emissions, since it has been observed in the atmosphere. Like chlorine, if it is emitted in sufficient quantities it can significantly enhance the photochemical reaction rates, though in this case no changes need to be made to the mechanisms to represent this. Currently NO_x emissions are assumed to include 2% of HONO, and CARB staff informs me that diagnostic calculations indicate that the model is not highly sensitive to HONO at this level. This may need to be verified with the updated models and emissions, particularly for low VOC/NO_x scenarios that will be most sensitive to radical initiators. HONO emissions from diesel vehicles may be much greater than that. This is indicated by environmental chamber studies with diesel exhaust in the Euphore chamber by Wiesen and co-workers, and also by chamber experiments in our laboratory. In particular, under CARB funding we did experiments with exhausts with a number of types of vehicles and fuels, and could get fair to good fits of model simulations to the experimental data for all except for the one experiment with diesel exhaust. Subsequent to writing the report I found that adding 5% HONO to the NO_x in the experiment. Work needs to be carried out to evaluate HONO emissions from diesel vehicles in California, and the sensitivity of the model to higher HONO from diesel emissions may need to be investigated.

Testing the VOC Inventory with Measurements

The model performance evaluation plan includes testing VOC emissions inventories by comparing speciated VOC emissions measurements with model predictions. This is done by converting the speciated VOC measurements to the corresponding lumped species that are represented in the model. The CARB staff recognizes the problems involved with spatial variations of primary pollutants, and presumably will adopt procedures to take this into account in the data analysis. But measurement biases are also important, as discussed below.

The main problem with speciated VOC measurements by GC is that they tend to be biased low, for several reasons. First, VOC emissions and oxidation products include polar and low volatility compounds that either can't be sampled quantitatively or don't make it through a GC column, or (usually) both. Appropriate sampling methods such as Tenax cartridges permit better sampling of such compounds, but this is not sufficient if the compounds can't make it through the GC column. Furthermore, almost all analyses include non-negligible amounts of unidentified GC peaks, and these are often not included or quantified in the analyses. Contractors conducting VOC measurement campaigns should always be required to report unknowns and provide estimates as to their quantities, but this practice has not always been followed in the past. Finally, multitudes of small peaks that are not resolved in the GC analysis can be treated as baseline using conventional analysis procedures. The significance of the latter is suggested by done with "2-D" GC analysis, where greatly enhanced resolution indicated that a substantial amount of the carbon is in small peaks that contribute to higher baselines and thus escape detection. The significance of undetected VOCs in general is supported by results of Paulson's work where a total carbon analyzer is found to detect more carbon than speciated GC analysis. Although this carbon is undetected, it may be represented in the emissions inventory and converted to model species in the simulations.

There are apparently no plans to test the VOC inventory using ambient total hydrocarbon measurements. Commercial FID total carbon analyzers are probably not useful because of the size of the methane interference in total carbon analyzers, because of the unreliability and problems with backflush models, and because of FID carbon response differences. However research into alternative total carbon analysis methods, such as that developed by Paulson, should be supported. Use of 2-D GC studies of air samples in the SCAB should also be investigated, if feasible. For SCOS evaluation, comparisons of total VOC measurements or 2-D GC analyses with speciated GC analyses using the methods employed during SCOS might provide a means to correct the data for inventory evaluation purposes. However, I suspect that current data are inadequate to assess whether the VOC inventory is high, but it may provide an indication whether it is low.

Evaluation Using NO_x and CO Data

Although CO is a non-negligible O_3 precursor, the primary importance of the CO inventory is that it provides a means to test meteorological and transport inputs to the model, because of its low consumption rate due to chemical reactions or deposition. Therefore, the accuracy of the CO inventory is important for that reason. Accuracy of the NO_x inventory is critical to O_3 modeling for reasons that are obvious to the CARB. Although NO_x species react rapidly, the consumption of total NO_x is slower, and comparison of modeled vs. measured NO_x in the morning in source areas also provides a test to the inventory or the relevant meteorological inputs. Because of the importance of NO_x emissions, any discrepancy between measured and modeled morning NO_x in the source areas would be a major concern. Comparisons of CO data with observations should be used to assess whether problem is due to meteorological or mixing model. If CO model agrees with data, possible biases in the NO_x inventory needs to be considered. NO_x and CO measurements on non-photochemical days could also be used to assess the inventory for these species, if relevant meteorological parameters are sufficiently well characterized.

Use of "NO_x" or "NO_y" data in non-source areas is *not* recommended for model performance evaluation except in the most qualitative sense. Fitz and co-workers in monitored a simplified irradiated VOC surrogate + NO_x air mass with multiple instruments used in the SCOS field study and found that different instruments of the same model give different readings. Data from the NO_x channel of converterbased analyzers should only be used for model performance evaluation if there is good reason to believe that the NO_x species present are primarily NO or NO_2.

Vehicle Emissions

The CARB recognizes the critical importance of having accurate vehicle emissions inputs in the models, and has undertaken major efforts to improve these inventories and remove the biases that have existed previously. This includes work in both traffic models and in vehicle emissions models. Some biases may remain, and because of the importance of vehicle emissions and the known biases in previous inventories, all known or suspected biases should be dealt with as discussed above in the "Treatment of Bias" section.

Some are concerned about whether the CARB's vehicle models adequately predict in-use emissions, and about possible biases that may result if deterioration factors are not correctly predicted. However, the CARB has an ongoing random in-use vehicle testing program that appears to go a long way towards providing the type of data needed. Reactivity of in-use emissions is also a concern, both evaporative and exhaust. The in-use vehicle testing program conducts speciation measurements on 10% of the in-use vehicles tested, both of the exhaust and the gasoline in the tank (the latter being useful not only for assessing evaporative speciation but also for the reactivity of gasoline in the marketplace). Presumably the results from this program are being incorporated into improving the vehicle emissions model and also the vehicle emissions speciation profiles.

Variability information (in speciation as well as mass) obtained in the vehicle emissions testing programs should also be used for uncertainty analysis. Because of its importance, systematic uncertainty analyses of vehicle emissions models should be carried out to assess the importance of this variability to model predictions, or at least to the total inventory.

There seems to be considerable ongoing work in improving vehicle activity estimates. The data for modeling freeway traffic activity appear to be adequate, but surface street activity is more uncertain, particularly for weekend differences as indicated above. The assumption that surface activity is proportional to freeway activity may be reasonable in the aggregate, though one would expect great variations in the weekday/weekend ratios relative to the freeway for surface streets in weekday-only business areas compared to residential neighborhoods or weekend-intensive retail areas such as by malls. Hopefully, weekend/weekday traffic differences are treated differently in these types of cases.

The reactivities of gasolines and exhausts appropriate for use in future-year scenarios are highly uncertain, but I am told they may be important in affecting attainment modeling. I am concerned about "reactivity neutral" assumptions that are apparently being made when deriving future vehicle emissions profiles. Gasoline formulators may well achieve future emissions standards using gasolines with varying reactivity, regardless of whether the standards are reactivity-based. If the reactivity of future exhaust is highly uncertain, it may need to be examined in diagnostic simulations. But artificial assumptions should not be made to make the uncertainty look like it is less than it really is, as it may introduce biases in the control strategy modeling. As discussed below, biases that are more important in the future-year scenarios than in the base case or evaluation simulations are of particular concern in this regard.

Biogenic inventory

The expanded domain used in the current modeling program may reduce the importance of background pollutants, but it makes biogenic emissions, and their corresponding uncertainties, relatively more important. Biases are also a particular concern in the biogenic inventory, because biogenics are expected to be much more important in the future year simulations than in the current episodes.

Diagnostic modeling already carried out has indicated that this is the case. This means that errors in the biogenic inventory may not have large effects on the performance evaluation or base case model results, but may significantly affect attainment demonstration or carrying capacity simulations. This would lead to biases in control strategy predictions.

The CARB recognizes the importance of reducing uncertainties in the biogenic inventories, and is supporting considerable research in this area. The various areas of research in this area were discussed in detail in the peer review meeting, and to a lesser extent in the modeling protocol document. The “GAP” database is used to estimate plant species distributions, work is ongoing to improve this, and the CARB Research Division has a project to look at these databases. The species estimates they incorporate are highly approximate but are better than nothing, and do not appear to be a source of known bias. The CARB participates in biogenic working group with EPA and others, but much of the work applicable to the East Coast not applicable to California, and vise-versa.

The one major concern I have with the current biogenic inventory concerns treatments of unknowns. Unidentified or unknown species constitute a non-negligible portion of the biogenic mass (one estimate I was given was ~30%), and yet they are not included in the biogenic databases. In other words, they are treated as if they are not there, even though they are really present and, like most biogenic VOCs, are probably quite reactive. This is a totally unacceptable and avoidable source of bias. Until data are concerning these unknowns are available, the modelers will need to contact appropriate estimates as to the amount of unknown mass that would remove this bias (see discussion of biases, above), and guidance as how best to represent them in the model. In the future, as indicated above, contractors need to be required to report the unknowns and provide estimates on their quantities and likely identities, because they are in a much better position to make more educated estimates than the modelers.

The BEGIS biogenic emissions database needs to be expanded to incorporate these “unknown” measurements or estimates and their recommended speciations. Note that consideration of most likely speciation of unknowns may vary with type of plant species, so the database should have the flexibility to add other types of chemicals.

In this regard, the speciation assignments for biogenic species should *always* be in terms of actual chemicals and *never* have direct assignments to lumped species used by the model. My understanding is that the biogenic speciation databases used by the CARB only allow for classification as isoprene, “terpene” or m-butenol. Note that different terpene isomers can have quite different reactivities and the SAPRC-99 is capable of representing these differences at least for the major isomers that have been studied. Therefore, the databases should allow for the ability to assign individual terpene isomers when such information is available. In addition, the databases should permit the assignments of unknowns to the most appropriate types of chemical species without constraint. The determination of how to represent these in the model is a chemical mechanism implementation issue that should not be mixed in with the inventory databases and assignments. This is discussed further below.

Speciation Databases and Emissions Speciation Processing

The CARB is making a significant effort to improve the quality of the speciation databases, though improvements are needed to the biogenic speciation database as indicated above. The most important are probably the vehicle emissions speciation databases, and work in this area is discussed above. The previously all-too-neglected stationary source inventory speciation profiles are also undergoing improvement, though this probably driven more by the CARB’s implementation or considerations of reactivity-based stationary source controls than any perceived importance of such profiles to SIP modeling.

Work is needed, however, to improve the organization of the speciation databases, and in particular the methods used to classify the various types of VOCs in the profiles. The categorization used by the CARB is no longer consistent with that used by the EPA, even though they were apparently the same at one time. The present categorization used by both includes complex mixtures and ambiguous classifications that are not real compounds, and assigns lumped model species directly to them. Directly assigning model species to mixtures is an amalgamation of chemistry and emissions that makes consistent treatment of the emissions by different mechanisms difficult, and complicates both new mechanism implementation and emissions speciation improvement. Mixtures should always be defined by specifying (or guessing) either the actual chemical compounds they are thought to contain, or by making a best estimate of a surrogate mixture of actual chemicals that might be an appropriate basis for representing them in a model. Model species should only be assigned to actual chemicals and never to mixtures. This allows mechanisms to be compared on the basis of the same emissions inventory assignments, frees the mechanism developer from having to make guesses about emissions, and frees emissions database developers from having to make judgements about particular mechanisms used in models.

The CARB is working on improving the speciation databases by gradually removing mixtures from its profiles as data become available. However, the EPA is probably not progressing as rapidly on this front, and neither the EPA nor the CARB are making any apparent attempt to make their categorization consistent. This inconsistency makes the problem of implementing new mechanisms into new databases almost twice as difficult. It is recommended that the CARB collaboratively support the development of a new speciation classification and mechanism assignment system that allows for consistent species classification and mechanism assignments in all regulatory models used throughout the United States. The system should be structured so that direct assignments of lumped model species to mixtures or profiles is no longer an option, but otherwise implemented so they can be interfaced to emissions processing systems[5].

Finally, it should be noted that my experience indicates that emissions processing procedures and databases provide the major obstacle to implementing new mechanisms in models or conducting systematic investigations of effects of mechanisms on model predictions. Since as discussed above comparing mechanisms should be an important part of the model evaluation process, work is needed to improve emissions processing procedures and databases to permit such evaluations to be readily conducted. For example, such improvements would make my recommendation to conduct diagnostic simulations using the RACM mechanism much more likely to be implemented.

Mexican Emissions

Uncertainties in emissions from Mexico can affect evaluation and control strategy simulations in San Diego and might also affect SCAB simulations as well, under certain transport conditions. The CARB recognizes the importance of this and is funding a contractor to conduct an independent review and update of a previous inventory developed by another contractor. However, future-year Mexican

[5] The reviewer has a conflict of interest in making this recommendation because he has distributed pre-proposals to do this work, as discussed in the “Reviewer Background” section above. Even if another contractor carried out this task, it would benefit the reviewer by making the job of implementing new mechanisms for the CARB and the EPA significantly easier.

emissions are particularly uncertain, and the problem of how to deal with this was not discussed in the protocol document or during the review meeting. Diagnostic simulations should be included to assess the effects of current Mexican emissions uncertainties on both the evaluation and the control strategy predictions, and of future Mexican emissions on the control strategy results.

I understand that the CARB is involved in a separate process involving border modeling issues, so these types of simulations may be planned as part of that process. However, some of the modeling for this process may be relevant to the ozone SIP, if only to show that Mexican emissions uncertainty is not a major issue in this regard.

Initial and Boundary Conditions

My impression is that for the most part there seem to be an adequate assessment of boundary condition uncertainties in the diagnostic simulations, with the possible exception of Mexico. No mention is made of transport from the east over the deserts or from Kern County. I assume this is because they have already been shown to be relatively unimportant or that they will be adequately evaluated in the diagnostic simulations. The possibility of very long range transport being influential in future year scenarios also needs to be assessed, as indicated above.

Model Evaluation Using Radical Measurements

I understand from Gail Tonnesen that models perform poorly in simulating radical levels measured in special studies. I understand that Bob O'Brien made some radical measurements in conjunction with SCOS97 but has no funding for data analysis. The CARB should consider including in the model evaluation process an analysis of this potentially valuable and unique dataset whose collection has already been paid for.

Technical Oversight and Review

The ozone SIP modeling is being carried out in coordination with a number of Southern California Stakeholder groups consisting of the affected regulatory or government agencies. They hold 3-4 meetings a year with these groups, and also have a separate meteorological working group including meteorologists from those agencies that meet periodically. Meetings of transport working groups, presumably also with various regulatory or government agencies, have begun, though the major effort in this regard is not yet underway.

This participation by stakeholder regulatory or government agencies is obviously necessary and important, and it appears that the CARB is doing what needs to be done in this regard. The UC Peer review provides a useful supplement to this, though the absence of an expert meteorologist among the reviewers is unfortunate, given the importance of meteorological issues in both the evaluation and policy planning process. Although the most technical reviewers will have some conflicts of interest and biases if they have sufficient expertise and experience to conduct a meaningful review, in general these will be different than those for most of the other stakeholders, and thus provides a useful supplement to the stakeholder input.

Another set of stakeholders that may not be adequately represented at this point are the various industry groups. Some of these significantly co-funded SCOS97, have considerable modeling and other relevant technical expertise, and presumably at are carrying out an independent analysis of the results. They will obviously provide input at the end of the process if the regulations affect their companies' bottom line. The CARB would be better off to have their input earlier in the process, when it is time to deal with it effectively.

The stakeholders should participate in the performance evaluation process and the considerations of the appropriate steps to take should performance be less than satisfactory. However, the CARB and the other regulatory stakeholder agencies are also under considerable pressure to have a satisfactory evaluation so they can go ahead with policy-relevant modeling that their leaders are really interested in. Therefore, the perception exists, however unfairly, that they may be tempted to gloss over or under-represent important problems revealed in the evaluation. For this reason, external objective peer review may be particularly useful during the period when the initial performance evaluation and associated diagnostic calculations are completed, but there is still time to do additional analysis before the policy recommendations have to be made. This is the time that decisions are made on what additional adjustments or diagnostics are needed, and how to appropriately proceed with the policy relevant analysis. External input at this point would probably be more helpful than at the end of the process, when the same criticisms would inevitably be raised, but when it will be too late to deal with them.

Summary Of Recommendations

I think the models and modeling procedures being proposed for use by the CARB for the 2003 ozone SIP for the most part represent the state of the art, and has significant improvements over past SIP modeling. I believe that the CARB staff understands the major issues and concerns involved, and are committed to conducting the most technically defensible modeling process possible for the resources available. However, the modeling process has many difficulties and uncertainties, and there will always be concerns about whatever process is employed and recommendations on how to do it differently. These concerns are listed below. Note that some if not most of these are probably being addressed at least to some extent in the present SIP or are priorities for future research, but are included in the list below either to emphasize their importance, or because I am uncertain of the CARB's priorities concerning them.

It is probably not feasible to carry out all of these recommendations for the 2003 ozone SIP, and some of the concerns and recommendations may already be adequately addressed. However, I believe that the CARB should consider the feasibility of carrying out at least some of these recommendations in the near term if they are not already being addressed, and in the other cases consider them as recommendations for future research.

- The portion of the documentation containing executive summary and recommendations for the policymakers is probably the most important output of this process. A major effort should be made to assure that it conveys all the caveats and uncertainties in a manner they can assimilate and properly interpret. The ability of the documentation to communicate the needed information to non-technical audiences should be evaluated using appropriate personnel as test subjects.
- Episodes modeled for control strategy assessment should represent the full distribution of meteorological conditions that are relevant to the problem being assessed, and not just those with sufficient data available for full model performance evaluation. Model performance evaluation should not be necessary for all episodes used for control strategy predictions, if the major components affecting such predictions are adequately tested in other representative episodes.
- Known biases should be removed from model inputs whenever possible. In some cases this can be done by adjustments so equal numbers of experts think they are high as low. Diagnostic calculations should be used to estimate effects on control strategy predictions of more complex biases that cannot be removed, such as grid resolution effects.
- Model components that are significantly out of date or have known biases or errors should not be

used except perhaps for diagnostic purposes if better alternatives are available. This is the case for the Carbon Bond mechanism, whose use in control strategy modeling should finally be discontinued. Condensed versions of up-to-date mechanisms should be developed if computational efficiency is the main reason that CB4 continues to be widely used.

- Models should never ignore unknowns that may have make non-negligible contributions, and must incorporate best estimates as to the amounts and nature, even if subjective guesses. Emissions databases need to include unknowns. This is particularly critical in biogenic emissions databases because ignoring biogenic unknowns can bias control strategy predictions.
- In this regard, contractors developing emissions inputs should be required to provide uncertainty and (if applicable) bias estimates for all the data they provide. They should also be required to report and estimate amounts of unknowns, and provide recommendations as to what types of compounds might be the least inappropriate to serve as the basis for representing them in models.
- Any adjustments or modifications made to the model specifically to improve its performance in simulating observed pollutant concentrations must be made with an appreciation that they may be covering up an error of a totally different type than the parameter being adjusted. All such adjustments should be clearly documented and diagnostic calculations should be used to evaluate the effects of alternative adjustments on control strategy predictions. Of particular concern is meteorological adjustments compensating for emissions or chemistry errors, or vise-versa.
- A second external peer review should be carried out after the results of the initial model performance calculations have been analyzed, but before final decisions have been made on how the control strategy modeling will be carried out. At least one of the reviewers should be an expert in meteorology who has not previously been part of this modeling process.
- Attempts should be made to obtain input from industry groups in the modeling process before it is too late to incorporate their input or effectively respond to their concerns.
- A priority should be given to determining whether the model can simulate observed weekend vs. weekday effects. This may require research into improving weekend inventories, particularly models for weekend traffic activity on surface streets.
- It should be feasible to systematically quantify effects on control strategy predictions for the major model components, such as emissions, that are based on scalar values. Subjective expert judgment can be used to obtain uncertainty ranges where needed. Diagnostic calculations should be used to estimate effects on control strategy predictions of more complex types of uncertainties, such as chemical mechanisms effects. Ideally, only one type of uncertainty should be varied at a time in such assessments.
- Because of mechanism uncertainties and possible errors in implementation, the CARB should not rely on only one mechanism for modeling. The RACM mechanism is about as up-to-date as SAPRC-99, is widely used in Europe, and may have a more accurate representation of some low NO_x reactions. It should be implemented in the CARB models to provide an up-to-date alternative to SAPRC-99 for evaluation.
- Software limitations that prevent useful diagnostic simulations from being carried out or that introduce known errors or otherwise unnecessary approximations into the model should be removed to the extent feasible. Chemical mechanism implementation and emissions processing software should be improved so that diagnostic simulations varying the chemical mechanism or lumping assignments can be more routinely carried out, and so that all the capabilities of SAPRC-99 for representing chemical detail can be incorporated. The quality, documentation,

flexibility, and ease of modification of model software should be among the criteria used in model selection.

- The effects of uncertainty in light intensity on evaluation and control strategy predictions need to be investigated. Some (but not all) of the light characterization data obtained during SCOS97 may be useful in assessing and reducing this uncertainty.
- The possibility that mathematical problems some models have in representing diffusion and transport may cause biases in model evaluations and control strategy predictions may need to be assessed, or at least ruled out.
- "Reactivity neutral" assumptions should not be made when estimating speciation profiles for future vehicle emissions, as they may introduce biases in control strategy modeling.
- The effects of uncertainties in Mexican emissions, particularly in future year scenarios, need to be evaluated.
- If ongoing research suggests that the role of chlorine needs to be investigated, then the CARB should take advantage of the major efforts underway in Texas concerning modeling chlorine chemistry.
- The possibility that HONO emissions from diesel may be significantly greater than from gasoline vehicles should be investigated.
- Evaluations of VOC emissions with ambient speciated measurements needs to appropriately take into the fact that such measurements are probably biased low. Research with improved total carbon analysis methods and very-high-resolution gas chromatography should be supported. Contractors conducting VOC measurement campaigns should always be required to report unknowns and provide estimates as to their quantities.
- Any discrepancies between observed and modeled NO_x and CO data in source areas should be analyzed to determine the extent to which they may indicate a problem with the NO_x emissions inventory, as opposed to uncertainties in meteorological inputs.
- The CARB should consider supporting analysis of radical measurements made during SCOS97 as part of the model evaluation process.
- Speciation assignments for complex mixtures in the anthropogenic inventory should always be in terms of actual chemical species and never as lumped species associated with particular chemical mechanisms. Contractors developing speciation profiles should be required to estimate what chemicals are most appropriate to represent unknowns or to serve as surrogates to represent complex or uncharacterized mixtures. The speciation databases and software used by the CARB and the EPA need to use consistent chemical classifications and not permit direct assignments of lumped model species to mixtures.
- "NO_x" or "NO_y" measurements made using commercial chemiluminescent analyzers should not be used for quantitative model evaluation. Eppley UV radiometer measurements should not be used for light model evaluation or input.

26 June 2001

John DaMassa
Planning & Technical Support Division
California Air Resources Board
P.O. Box 2815
Sacramento, CA 95812

Dear John,

Attached please find my detailed review comments on the modeling protocol document for southern California that you provided. A summary of my major comments is provided below. This review was completed under interagency agreement 98-004 between the University of California and the California Air Resources Board.

The protocol document is an important communication mechanism within the modeling team, and with external stakeholders and technical experts. A clearer statement of objectives for the modeling study is needed at the beginning of the protocol document.

The most adverse meteorological conditions leading to high ozone levels were not seen during SCOS97; it is appropriate to consider other more severely polluted episodes as part of an attainment demonstration for southern California.

Major reductions in air pollutant emissions and ozone levels were observed between 1987 and 1997 in southern California. I recommend updated emission estimates be developed for a summer 1987 (SCAQS) episode, and modeling be conducted of the changes in ozone that took place between 1987 and 1997.

Numerical guidelines for model performance, measured in terms of normalized bias and gross error of the ozone predictions, have been overemphasized. Decisions about which models and input data are “best” and “state-of-science” should not focus only on agreement between observed and predicted ozone levels. Many compensating errors may be present in model algorithms and inputs, and better models may reveal problems that were previously hidden.

Attempts should be made to understand and reconcile potentially large and influential uncertainties in off-road diesel engine NO_x emissions.

This modeling study continues California's tradition of being at the leading edge of developing model inputs and using photochemical models to develop air quality management plans. I am very impressed by all the progress that has been made on many fronts since SCAQS in 1987. I encourage you to publish in peer-reviewed journals some of the insights that will surely be gained in this effort.

I hope my comments are constructive and will be helpful to you as the modeling effort continues. Please call me at (510) 643-9168 if you have any questions.

Sincerely,

Robert Harley
Associate Professor

Attachment: detailed peer review comments.

Review of

Modeling Protocol for Regional 1-Hour and 8-Hour Ozone Modeling in Southern California for the 2003 State Implementation Plans

Inter-Agency Agreement 98-004

Prepared for

Planning & Technical Support Division
California Air Resources Board
P.O. Box 2815
Sacramento, CA 95812

Prepared by

Robert Harley, Associate Professor
Department of Civil and Environmental Engineering
University of California
Berkeley, CA 94720-1710

June 26, 2001

Introduction

The subject of this peer review is a photochemical modeling protocol document for southern California, draft #3 dated September 18, 2000. In addition to reviewing the protocol document, the review process included a daylong meeting with ARB and SCAQMD staff in Sacramento, and a follow-up conference call on mobile source emissions issues. Although other University of California scientists are expected to be involved in reviewing the modeling protocol, the review comments presented here were developed without consulting other reviewers. This approach was desired by ARB staff to provide a greater range of peer review comments, and more opportunity for staff to interact with individual reviewers.

According to DaMassa (1992), the modeling protocol process and document are essential features of a modeling study, laying the foundation for the technical work that will follow, and providing an early and appropriate opportunity for external review before the work is conducted.

Background and Objectives of Modeling Study

Major field studies of photochemical air pollution were conducted in southern California in summer 1987 (Southern California Air Quality Study or SCAQS) and summer 1997 (Southern California Ozone Study, SCOS97). The protocol document provides an appropriate overview of SCOS97, including objectives and scope of the field study. *The objectives of the proposed photochemical modeling work should be listed and described more clearly at the beginning of the protocol document*. For example, modeling objectives could include some or all of the following: (1) model evaluation for recent (1997-98) air pollution episodes in southern California, (2) assessment of inter-basin transport of ozone and precursors, (3) selection of state-of-science model(s) to use in future air quality planning efforts, (4) assessment of changes in precursor emissions and ozone levels in southern California between 1987 and 1997, (5) attainment demonstration for ozone in the South Coast Air Basin and other neighboring areas, (6) study of weekday-weekend differences in ambient ozone levels and precursor emissions. There may be other objectives as well.

Modeling Domain

The proposed geographic extent of the air quality modeling region is presented in Figure 1 of the protocol document. This domain is large, extending well beyond the urbanized areas of the South Coast Air Basin. Such a domain is necessary for photochemical modeling if inter-basin transport issues are to be assessed.

On page 3, the statement that “terrain elevations in Southern California often exceed 2000 m AGL” is incorrect, because local terrain height is always defined to be 0 m above ground level (AGL). I think the authors mean 2000 m MSL (i.e., above mean sea level).

Choices of different locations for boundaries of the modeling domain will not necessarily minimize the influence of boundary conditions on simulation results… the numerical values of the boundary conditions used in the modeling are also important.

Episode Selection

Concerns about past regulatory photochemical modeling studies in southern California included a lack of weekend episodes (when ozone concentrations tend to be higher on average), and dropping a severe episode from summer 1985 from the attainment demonstration in the 1997 AQMP.

Six different air pollution episodes from 1997 and 1998 are proposed for further consideration in Table 1. The list includes a weekend episode (26-29 September 1997, a Friday through Monday), and several other episodes that include one or more weekend days. Unfortunately, SCOS97 failed to capture any high-ozone episodes describing photochemical air quality in the basin under the most adverse meteorological conditions. *It would be helpful to include in the protocol document a review of the El Niño effect*, and discuss how it may have affected precipitation, cloudiness, and temperature in southern California in 1997 and 1998.

Whereas the episodes from 1997 have an extensive supporting aerometric database from SCOS97, the 14-19 July 1998 episode listed in Table 1 was more severe than any of the others listed. I*t would be appropriate to consider the 14-19 July 1998 episode as part of a determination of the emission-carrying capacity of the South Coast Air Basin and neighboring areas*.

The summer 1987 SCAQS air pollution episodes remain of historical interest. *Inclusion of one SCAQS episode (e.g., 23-25 June 1987) should be considered*, as part of an assessment of how emissions and air quality have changed in southern California over the intervening decade between 1987 and 1997.

Air Quality Model Selection

The modeling protocol mentions concerns that the current regulatory air quality model (UAM-IV) has limitations that make it unsuitable for regional-scale photochemical modeling. A variety of newer models including CALGRID, CAMx, CMAQ, SAQM, UAM-FCM, and UAM-V are mentioned as candidate models for further study.

I recommend against selecting or deselecting models based solely on the level of agreement between observed and predicted ozone levels. The possibilities for compensating errors in model inputs and model algorithms are legion. A better model may appear to be inferior in terms of ozone performance, only because an improved description of one or more processes may reveal other problems that were previously hidden. Detailed review of model components such as advection algorithm, chemical mechanism, dry deposition module, numerical integration scheme, documentation, ease of input data preparation and output postprocessing, etc. should all be considered as factors to decide which models are to be recommended for use.

At the bottom of page 9, the text states that ARB and SCAQMD staff have agreed to begin testing UAM-V, among other models, for the SCOS97 domain. This seems inconsistent with text immediately above stating that the present non-public domain status of UAM-V precludes regulatory usage.

I am concerned about the multiplicity of air quality model/air pollution episode combinations. There are 6 different air quality models, and 6 or more air pollution episodes proposed for further study. *Applying all of the proposed air quality models to all of the air pollution episodes is unlikely to be a productive use of staff time*. A strategy for deciding on which air quality model(s) will be used should be described, and this strategy should be executed early so the multiplier effect doesn't overwhelm the ability to gain and communicate insights from the modeling effort.

Chemical Mechanism

In the past, air quality models contained a "hardwired" chemical mechanism that was an integral part of the model and could not easily be changed. More recently, air quality models have evolved so that the chemical mechanism can often be described as a model input. The Carbon Bond-IV mechanism (Gery et al., 1989) lacks options for detailed resolution of VOC chemistry, and is no longer "state-of-the-science" in my opinion.

VOC lumping strategies should be discussed in the protocol document. It may be worthwhile to include MTBE and ethanol as explicit species in the chemical mechanism, since major changes in emissions of these species have occurred or will occur in southern California due to gasoline reformulation. It may also be worthwhile to include methylbutenol (MBO) as a separate, explicit model species in the mechanism, since the biogenic emission inventory now includes this species. In general, biogenic and anthropogenic VOC should not be lumped together, since their spatial and temporal patterns differ, and no single characterization of lumped species properties will apply throughout the modeling domain if biogenic and anthropogenic VOC are combined. These changes should help stabilize the mix of VOC assigned to each lumped species, and hopefully minimize changes in lumped species properties over time.

In summary, *more detailed discussion of chemical mechanisms and VOC lumping should be added to the protocol document.*

Horizontal and Vertical Grid Resolution

The horizontal grid resolution is proposed to be 5 km, which is consistent with past practice. The extent of the modeling domain is larger now, so there will be more grid squares. MM5 and CALMET provide wind fields and other meteorological inputs on different grid systems – Lambert Conformal for MM5 and a UTM-based grid system for CALMET, so two different sets of gridded emission estimates will be needed.

To facilitate inter-comparison of model predictions, *it is recommended that the same or similar vertical cell heights be used where possible*, especially for the lowest layer near the earth's surface. Large differences may arise in primary pollutant concentrations at night if pollutants remain trapped within the ground-level cell. Tables 2 and 3 currently indicate different ground-level cell heights of 61 and 20 meters, respectively. In Table 2, it is not clear why the air quality model extends all the way to the top of troposphere. Is it possible for significant exchange to occur between the uppermost cells and lower layers on the time-scale that air masses remain within the modeling domain? In Table 3, the vertical heights for cells 4 and 5 are incorrect.

Meteorological Inputs

The protocol document discusses three different types of models that can be used to generate the meteorological inputs needed for air quality modeling. The objective and diagnostic models rely on interpolation of observations, whereas prognostic models solve coupled conservation equations for fluid mass, momentum, energy, and water content. A procedure known as four-dimensional data assimilation (FDDA) allows the predictions of prognostic models to be nudged towards observations when and where they are available. A more detailed review is provided by Seaman (2000).

The southern California modeling domain includes complex topography, with mountains as high as 3000 meters or more above sea level, with associated upslope and downslope flows that are difficult to model. Meteorological observations are sparse within mountainous and unpopulated desert areas. Inter-basin transport sometimes occurs through mountain passes (e.g. Banning Pass connects the South Coast and Salton Sea air basins). Development of accurate meteorological fields for this domain is a challenging task, regardless of the modeling approach that is used. The larger modeling domain and the desire to assess inter-basin transport place greater demands on the met modeling, especially with respect to flow over complex terrain.

The use of radar wind profilers has greatly expanded the frequency of wind observations aloft, though relative to SCAQS we may have lost ground in obtaining temperature and humidity data aloft (soundings conducted using balloons typically provide vertical temperature and humidity profiles that are useful in locating the inversion layer). Further consideration of the adequacy of the observational database (wind, temperature, and humidity) throughout the full horizontal and vertical extent of the modeling domain is needed before objective/diagnostic modeling approaches can be considered as a source of met inputs for the air quality models.

On page 16, there is a vague statement that met-fields "will be evaluated to determine which is most suitable for air quality modeling." The text beginning on page 17 (validation and technical review) provides more specific methods for evaluation. The ability to simulate tracer release experiments accurately could be helpful in making decisions about which met fields to use in air quality modeling. *Decisions about which met-fields will be used should not be made by looking for best agreement with observed* *ozone levels*, because of the possibilities of compensating errors in other model inputs and algorithms, as discussed above.

In the past, separate met modeling approaches have been used for individual episodes within the scope of a single planning study. This situation should be avoided in the future. However, it may be necessary to use different met modeling procedures for the 14-19 July 1998 episode, since not all of the met observations from SCOS97 continued in the following summer.

Emission Inventory

Assembling accurate emission estimates throughout the large southern California modeling domain will be a very challenging task. There are many different agencies involved in the effort, including numerous local air pollution control and transportation/ land use planning agencies, Air Resources Board staff in El Monte and Sacramento, and various contractors.

Point and Area Sources

Point source emission estimates are available for 1996 and are being compiled for 1997. An emission forecasting system (Johnson and Lakhanpal) will be used where point source estimates are not available for the year and sub-region in question. The situation is similar for area sources. Updating the inventory is difficult because of the need to reconcile point and area source estimates to prevent double-counting (Bickett). Furthermore, emission estimates must be collected from local air districts and ARB staff.

Mobile Sources

Concerns about past photochemical modeling studies in southern California focused on questions of negative bias in the hydrocarbon emission inventory, representation of weekday vs. weekend changes in motor vehicle activity and emissions, and representation of the different spatial and temporal patterns appropriate for heavy-duty diesel trucks. Major efforts have been made by ARB staff over the last 10 years to address these and other concerns. New models (EMFAC 2000 and OFFROAD) have been developed to estimate mobile source emissions, and hydrocarbon emission estimates have increased compared to earlier versions of EMFAC. Emission sources such as leaking liquid gasoline from poorly-maintained vehicles have been recognized and are now estimated in the emission inventory. A freight-based travel demand model and weigh-in-motion traffic counters are being used to improve the representation of diesel truck travel and emissions in southern California.

ARB staff have used a variety of methods to assess estimates of mobile source emissions (ARB, 2001). These verification methods include comparisons with Van Nuys tunnel results from 1987 and 1995, comparison with remote sensing-based estimates of CO and HC emissions obtained during SCOS97 (Singer et al., 2000), and comparisons of fuel consumption estimates from the models with statewide fuel tax returns. It is surprising that EMFAC 2000 is predicting no change in diesel fuel consumption in the state through the 1990s.

The Van Nuys tunnel provided unique and early on-road measurements of vehicle emissions from 1987, but this tunnel is short (220 meters) and uncertainties in measured emission factors tend to be large. For more recent years (1995 and 1996), measurements by Gertler et al. (1999) at the longer Sepulveda Tunnel in southern California may provide more useful points of reference for comparison with EMFAC model predictions.

Other assessments that should be reviewed address questions such as the relative importance of tailpipe vs. evaporative sources of HC emissions (Pierson et al., 1999), gasoline vs. diesel engine contributions to on-road vehicle NO_x emissions (Kirchstetter et al., 1999; Sawyer et al., 2000), the importance of off-road diesel engines as a source of NO_x emissions (Kean et al., 2000), and the contributions of cold start vs. stabilized operating conditions to total exhaust emissions (Singer et al., 1999).

A potentially large and influential uncertainty is off-road diesel engine activity and associated NO_x emissions. Kean et al. (2000) used fuel sales surveys conducted by the U.S. Energy Information Administration (EIA) to estimate off-road diesel engine activity. In contrast, the OFFROAD model relies on estimates of number of engines, hours of use, and load factor, broken down by horsepower rating and end-use application of the engines. Using statewide data for California from 1996, off-road diesel NO_x emissions in the construction and agricultural sectors are estimated to be 98 and 37 tons per day, respectively (see Kean et al. 2000 for supporting data). California's OFFROAD model estimates for calendar year 2000 predict NO_x emissions that are ~3 times higher. Differences in activity estimates were identified by Kean et al. as a source of uncertainty. *Uncertainties in the estimates of off-road diesel engine emissions are large enough to recommend further in-depth study, and consideration of alternate emission scenarios in the photochemical modeling if the estimates cannot be reconciled.*

Major efforts to control mobile source emissions were made through the 1990s, including more effective and durable control technologies installed on new vehicles, reformulated fuels, and enhanced inspection and maintenance. A dramatic decrease in photochemical air pollution in southern California has been observed over the same time period. *Historical emission estimates for 1987 should be developed to support modeling of the changes in ozone levels that have occurred between SCAQS in 1987 and SCOS97.*

Day-Specific Emissions

Changes in environmental conditions (mainly temperature) and day of week (especially weekday vs. weekend) lead to changes in air pollution emissions. Unusual events such as wildfires and shutdowns at major point sources can also affect emissions and air quality. For aircraft, it would be useful to develop day-of-week estimates (typical weekday, Friday, Saturday, and Sunday). The differences from day to day should be assessed, to provide guidance as to whether data should be collected for each specific modeling episode, or whether generic day-of-week estimates can be used.

Natural Emissions

Emissions of isoprene, monoterpenes and methylbutenol will be estimated for the modeling region using a biogenic emission inventory system, BEIGIS. *As anthropogenic emissions decline, the relative importance of biogenic emissions is likely to increase.* Compared to estimates used in past modeling, updates to the biogenic emission inventory include new plant species data as a function of land use, updated emission factors and dependence on environmental conditions, and inclusion of a new species (methylbutenol) emitted by pine trees.

Organic Gas Speciation

Reformulation of gasoline in California that took effect in 1995-96 has led to major changes in VOC speciation from mobile sources. New speciation profiles are therefore needed to represent conditions in 1997/1998.

Northern Mexico Inventory

It is not clear whether the modeling domain shown in Figure 1 includes enough of northern Mexico to make a meaningful assessment of air pollution transport between Mexico and southern California. How will the boundary condition at the southern edge of the modeling domain be specified?

Initial and Boundary Conditions

The modeling region is larger than used in past studies, so the "spin-up" time required to remove the influence of uncertainties in initial conditions may be longer now. *It is recommended that simulation spin-up should begin during the mid-afternoon rather than at midnight*, so that surface observations used to specify IC are more likely to represent pollutant concentrations in the first ~500 meters accurately. Using observations from midnight to initialize the model is likely to represent concentrations in only the first ~50 meters of the modeling domain.

As noted above, *prescribing a large modeling region does not guarantee the influence of boundary conditions is minor*. The numerical values used for the boundary conditions are also important. The amounts of formaldehyde, ozone, PAN, and NO_x at the boundaries may be particularly influential. The rationale and methodology for altering BCs for modeling future-year scenarios should be presented. Sensitivity studies are needed to assess the importance of BCs, as described on p. 31 of the protocol document.

Model Performance Evaluation

Numerical guidelines for model performance, measured in terms of normalized bias and gross error (DaMassa, 1992) are a useful point of reference, but *have been over-*

emphasized, and may lead to bad decisions about the most appropriate air quality models and input data to be used.

In addition to evaluating model predictions for ozone using graphical and statistical techniques, and investigating model responses in a series of diagnostic simulations, I *recommend evaluations of model predictions for total NMOC, nitric oxide* (NO is measured directly by chemiluminescent analyzers, CLA), *and* NO_y (meaningful only at sites with externally mounted converters on the CLA). Comparing the distribution of total NO_y among NO, NO_2, PAN, PPN, HNO_3, alkyl nitrates, etc. would be another useful test of model predictions vs. observations, though there are probably only a few research sites where this comparison of NO_y speciation is possible. Evaluation of chain-initiating (HCHO, HONO) and chain-terminating (H_2O_2, CH_3OOH, and HNO_3) species predictions may also be possible at a few sites.

Use of the Modeling Results

Applications for the modeling results include estimation of the carrying capacity for ozone precursors, demonstration of attainment of air quality standards, and assessment of inter-basin transport. *A useful test case for evaluating the relative response factor (RRF) would be to look at changes that occurred in observed and predicted air quality in southern California between 1987 SCAQS and SCOS97*.

The protocol document should clarify the rationale for looking at 1-hour peak ozone concentrations, now that EPA's 8-hour ozone standard has been sustained in court and the Federal 1-hour ozone standard has been abolished.

Deliverables

The protocol document envisions a large effort involving many people and a year or more of effort. The regulatory mission of the agencies involved means that a top priority must be development of state implementation plans for attainment of air quality standards. If time and resources permit, *it would also be desirable to publish selected aspects of this modeling study in the peer-reviewed literature*. This would provide appropriate recognition and visibility for staff efforts, and a means for continuing professional development. Rather than a single omnibus paper, more focused reports could be written with individual staff members taking the lead in presenting noteworthy findings.

References

ARB (2001). SB 2174 Verification Methods. Personal communication from Jeff Long, Mobile Source Control Division, California Air Resources Board, El Monte, CA.

Bickett, C.A. Redesign of the California emission inventory computer system. Technical Support Division, California Air Resources Board, Sacramento, CA.

DaMassa (1992). Technical Guidance Document for Photochemical Modeling. Technical Support Division, California Air Resources Board, Sacramento, CA.

Gertler, A.W.; Sagebiel, J.C.; Dippel, W.A.; O'Connor, C.M. (1999). The impact of California phase 2 reformulated gasoline on real-world vehicle emissions. *Journal of the Air & Waste Management Association* **49**, 1339-1346.

Johnson, M.E.; Lakhanpal, B. Redesign of California's emission forecasting system. Technical Support Division, California Air Resources Board, Sacramento, CA, and California State University, Fullerton, CA.

Kean, A.J.; Sawyer, R.F.; Harley, R.A. (2000). A fuel-based assessment of off-road diesel engine emissions. *Journal of the Air & Waste Management Association* **50**, 1929-1939.

Kirchstetter, T.W.; Harley, R.A.; Kreisberg, N.M.; Stolzenburg, M.R.; Hering, S.V. (1999). On-road measurement of fine particle and nitrogen oxide emissions from light- and heavy-duty motor vehicles. *Atmospheric Environment* **33**, 2955-2968.

Pierson, W.R. et al. (1999). Assessment of nontailpipe hydrocarbon emissions from motor vehicles. *Journal of the Air & Waste Management Association* **49**, 498-519.

Sawyer, R.F.; Harley, R.A.; Cadle, S.H.; Norbeck, J.M.; Slott, R.; Bravo, H.A. (2000). Mobile sources critical review. *Atmospheric Environment* **34**, 2161-2181.

Seamen, N.L. (2000). Meteorological modeling for air-quality assessments. *Atmospheric Environment* **34**, 2231-2259.

Singer, B.C.; Kirchstetter, T.W.; Harley, R.A.; Kendall, G.R.; Hesson, J.M. (1999). A fuel-based approach to estimating motor vehicle cold start emissions. *Journal of the Air & Waste Management Association* **49**, 125-135.

Singer, B.C.; Harley, R.A. (2000). A fuel-based inventory of motor vehicle exhaust emissions in the Los Angeles area during summer 1997. *Atmospheric Environment* **34**, 1783-1795.